Wind and solar's biggest hurdle: hooking to the power grid p. 1042

Early detection of transplant rejection pp. 1048 & 1105

Cerebellar granule cells remodeling across the life span pp. 1049 & 1112

Science

$15
8 SEPTEMBER 2023
science.org

AAAS



FROM TRASH TO TREES

Nudging green choices by forgoing single-use cutlery in China's food delivery apps p. 1064

VISIT…

LANZAROTE Caliente.com

EDITORIAL

Embed equity throughout innovation

The social benefit of technologies is frequently unevenly realized across the United States. Rural communities, individuals with disabilities, and historically marginalized groups face out-of-reach costs or lack access to products that meet their needs. Blame is typically placed on complicated regulatory processes or complex delivery systems, but this response neglects the problem that equity is not baked into the nation's innovation process at any stage. The United States needs to rethink its entire innovation ecosystem to incorporate equity as a foundational guiding principle—from research design and funding requirements to policies and regulations that govern the delivery and oversight of new products to the public.

Development of new technologies and products in the United States benefits from a governance framework that optimizes fairness and opportunity for creators and investors while prioritizing the safety and autonomy of end users. Equity is typically an afterthought, usually arising after unfair public outcomes are recognized. In health care, for instance, recent examples include the restrictive cost of new drugs to treat hepatitis C and the COVID-19 diagnostic test kits that were mass produced but didn't reach marginalized communities. Policies that foster innovation have focused on profitable transitions of scientific breakthroughs from discovery to the market. Particularly in the decades after World War II, this approach fueled innovation and substantially bolstered the nation's economy. However, a primary aspiration of innovation—public good for all—has often been unrealized.

"...equity is not baked into the nation's innovation process..."

Given the recent remarkable pace of progress across all scientific disciplines, there is an urgent need to incorporate equity considerations throughout the innovation life cycle. Consider, for example, the increasing use of artificial intelligence (AI) and machine learning in health care. Lack of diversity in the workforce and in the geographic distribution of innovation centers reduces the range of perspectives that shape research in this field. Nonrepresentative and discriminatory data sets and poor selection of proxy outcomes measures lead to biased algorithms that skew machine learning and adversely affect AI platforms that influence decision-making. Although efforts are underway to advance the ethical use of AI, a fundamental cultural shift is needed to fully integrate equity in this field and in many others.

Embedding these values in innovation means accounting for equity at each stage, from idea conception to technology and product development, evaluation, monitoring, and iterative learning and improvement. In its recent report, *Toward Equitable Innovation in Health and Medicine: A Framework*, the US National Academies of Sciences, Engineering, and Medicine and the US National Academy of Medicine called for a systems approach to implement equity-promoting practices. The recommendations include using policy and regulatory levers to set funding priorities; developing new metrics and equity-based models of technology assessment; and providing federal support for the development of methods, metrics, and benchmarks for achieving equity.

The report also emphasizes that a diverse community of users and creators should join traditional stakeholders to rethink governance strategies and incentives. Equitable innovation will require creative steps and new practices, such as engaging with underserved and marginalized communities at each stage of the innovation life cycle. This will ensure that these groups are not only consulted during research but also have opportunities to substantively lead and engage in innovation partnerships.

Addressing equity in innovation is evident in recent federal initiatives. An executive order issued in 2023, Further Advancing Racial Equity and Support for Underserved Communities Through the Federal Government (following an earlier one in 2021), calls on agencies to assess gaps and opportunities to enhance equity across the entire government; action plans were released in 2022, with annual updates required. The Office of Science and Technology Policy (OSTP) also has recognized the problem, and in 2022 it released a vision for equity in the US science and technology ecosystem that emphasizes support for students, teachers, workers, and historically underrepresented communities. These are crucial first steps, but there is a strong need for a coordinating body—an OSTP-helmed Equity in Biomedical Innovation Task Force—to translate federal priorities into goals and actions that span all facets of innovation, from funding choices to research design to regulatory policies.

If 21st-century innovation is to reflect social needs and improve the well-being of the entire public, then it will require a strong vision activated by a coalition of public and private partners that embrace an equity-centric approach at every stage.

–Keith A. Wailoo, Victor J. Dzau, Keith R. Yamamoto

Keith A. Wailoo is Henry Putnam University Professor of History and Public Affairs at Princeton University, Princeton, NJ, USA. kwailoo@princeton.edu

Victor J. Dzau is president of the National Academy of Medicine, Washington, DC, USA. vdzau@nas.edu

Keith R. Yamamoto is vice chancellor for Science Policy and Strategy and director of Precision Medicine at the University of California, San Francisco, CA, USA. He is president of the American Association for the Advancement of Science (the publisher of *Science*), Washington, DC, USA. yamamoto@ucsf.edu

10.1126/science.adk6365

NEWS

> "The worry is that the first time that it goes through, it will really have a high impact in … mortality."

Veterinary pathologist Thijs Kuiken, in *The New York Times*, on forecasts that the H5N1 avian influenza virus will soon ravage previously unexposed bird and mammal populations of Antarctica.

IN BRIEF

Edited by **Jeffrey Brainard**



Water hyacinth, an invasive species and the most widespread terrestrial weed, harms fisheries and impedes boats.

ECOLOGY

Annual cost of invasive species put at half-a-trillion dollars

Invasive species cause more than $423 billion per year in damage to agriculture, fisheries, water supplies, and other ecosystem-dependent benefits worldwide, according to the summary of a comprehensive review by dozens of scientists, released this week. The monetary losses, adjusted for inflation, have quadrupled every decade since 1970, the study's baseline, the summary says. The report is the first on the topic from the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, which has 143 member nations. The estimated financial loss is "a huge, huge underestimate," Helen Roy of the UK Centre for Ecology & Hydrology, who co-chaired the group that wrote the report, said in a media briefing; many costs such as weeding invasive plants have not been quantified, she said. More than 3500 species are known to have become invasive after people moved them, intentionally or unintentionally, to new locations where they have crowded out native plants and animals, some of which supported local economies. The number of invasive species is rising faster than ever because increases in global trade and travel help spread them, the summary says. But only 17% of countries have laws or regulations to prevent or manage invasions of these species.

South Korea eyes research cuts

FUNDING | South Korea's government has proposed cutting the country's research budget for 2024 by 11%, to reduce inefficiencies and the number of low-priority projects. Funding for basic research will be cut 6% and for national research institutes, 9%; the total for science and engineering will be 25.9 trillion won ($19.5 billion), the Ministry of Science and ICT said on 30 August. The plan increases support for younger researchers and fields such as artificial intelligence and biomedicine. In recent years, increasing governmental and private sector investment has raised South Korea's R&D spending to 5% of its gross domestic product, trailing only Israel's 5.9%, according to the Organisation for Economic Co-operation and Development. In comparison, the United States devotes

just 2.6%. Some researchers say the government could improve evaluations of research programs but warn the cuts to basic work will undercut the nation's scientific prowess. The National Assembly must still approve the budget.

Cancer rises in younger adults

GLOBAL HEALTH | Cancer cases in adults under age 50 worldwide have increased nearly 80% in the past 30 years, while deaths have grown by more than 25%, a study has found. When adjusted for changes in the global population, the results point to a strong rise in cancer rates but a decrease in mortality over time. The study, published this week in *BMJ Oncology*, is one of the largest of its kind and draws from data collected by the Global Burden of Disease 2019. The burden of disability and death among people ages 14 to 49 was greatest for colorectal cancer followed by other types—breast, lung, and stomach cancers—less commonly associated with younger adults. The researchers cite risk factors such as alcohol and tobacco use, physical inactivity, and excess weight, factors also associated with cancer among older adults. They propose that early screening and prevention programs to improve lifestyle and diet could help, particularly among the most affected group, people ages 40 to 49.

Distant galaxy had magnetic field

ASTRONOMY | Scientists have spied a magnetized galaxy that existed 11 billion years ago—the earliest such body detected to date. The Milky Way and nearby galaxies have magnetic fields that help clouds of gas collapse into newborn stars. But researchers didn't know when these fields emerged. This week in *Nature*, astronomers report using the Atacama Large Millimeter/submillimeter Array—a radio observatory in Chile—to observe a distant galaxy called 9io9, which existed when the universe was less than one-fifth of its current age. Its light is slightly polarized, evidence that a magnetic field aligned dust grains in the galaxy, which then emitted photons that vibrate in a favored direction. Additional research may illuminate the role played by distant magnetic fields in the furious pace of star creation in the young universe.

Strike over outbreak alerts ends

PUBLIC HEALTH | The staffers who run ProMED, an early warning email system about infectious diseases, plan on 11 September to end their monthlong strike against the scientific society that runs the free service. Public health specialists and journalists have relied on the service for the latest reports on outbreaks of diseases such as COVID-19 and Ebola. The strike followed a decision in July by the International Society for Infectious Diseases (ISID), which operates the service, to require paid subscriptions, citing a lack of sustained funding. The strikers, who comprise the service's moderators and editors, demanded that ISID seek financial partners, ensure their editorial independence, and improve transparency. ISID told the strikers it is committed to finding financial partners and wants to improve communications and establish an advisory group to help with decisions. As a "gesture of good faith," the strikers told ISID in an email, they agreed to go back to work while it works on "definitive solutions."



PHOTO: PATRICK SEMANSKY/AP PHOTO

BIOSAFETY

White House mulls refining pathogen rules

The White House last week requested public comment on whether to adopt controversial recommendations that the government tighten oversight of federally funded research on dangerous pathogens. In March, the National Science Advisory Board for Biosecurity (NSABB) called on the government to expand the swath of research subject to special reviews and stricter regulation. Targets included gain-of-function studies that could give pathogens new abilities to infect humans, and dual-use research that could be used for good or harm. NSABB's recommendations have drawn protests from some microbiologists, who fear tighter restrictions could hamper important, low-risk disease research. On 1 September, the White House Office of Science and Technology Policy (OSTP) asked for input on a range of issues, including how to identify studies that could be "reasonably anticipated" to pose unacceptable risks and whether to narrow NSABB's recommended policies. OSTP is seeking responses by 16 October.

A U.S. panel proposed tighter rules for pathogen research, some of which requires biosafety suits.

Paris nouveau: insecticide fog

PUBLIC HEALTH | Paris last week sprayed insecticides to kill the *Aedes albopictus* mosquito that spreads dengue, in the city's first mosquito fumigation campaign. The trigger for the spraying was two cases of dengue in people who reportedly developed the disease shortly after having traveled outside the country. The European Centre for Disease Prevention and Control says France also had many locally acquired cases—65 in 2022 and four this year, most in the southern part of the country that has the Mediterranean climate this mosquito species prefers. Neighboring Spain and Italy have also reported locally acquired cases over the past 5 years. France first detected *A. albopictus*—also known as the Asian tiger mosquito—in 1999, but it died out, only becoming endemic after a new introduction in 2004. The species, which can also transmit chikungunya and Zika virus, may be heading northward in Europe because of climate change.



In 2019, the firm Natural Hydrogen Energy drilled the first U.S. well in search of geological hydrogen.

IN DEPTH

ENERGY

Geological hydrogen wins first major funding

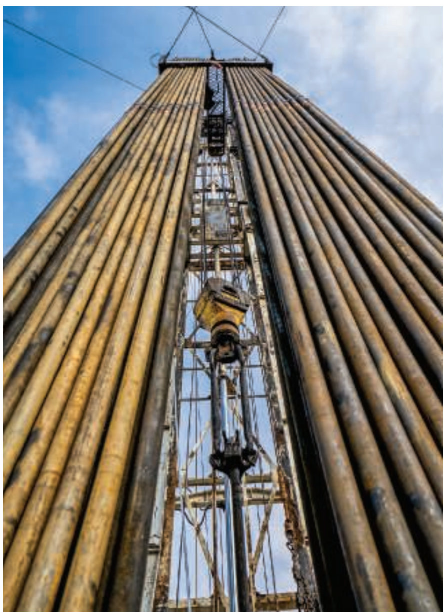
$20 million grant program by ARPA-E aims to jump-start work on climate-friendly fuel

By **Eric Hand**

A dark horse concept in the race to develop clean and sustainable energy sources is getting its first major investment from the U.S. government. This week, the Advanced Research Projects Agency-Energy (ARPA-E), the high-risk, high-reward arm of the Department of Energy (DOE), announced it would fund $20 million in grants to advance technologies for extracting clean-burning hydrogen from deep rocks. At the moment, all of the world's hydrogen is manufactured industrially. But some researchers have concluded that, contrary to conventional wisdom, Earth harbors vast deposits of the gas that could be tapped like oil—and that reserves could be stimulated by pumping water and catalysts into the crust.

The ARPA-E funding "will be the largest single investment in R&D of this nature worldwide," says Yaoguo Li, a geophysicist at the Colorado School of Mines who plans to apply to the program. "It's so new, a lot of other agencies and other countries are just waking up to this possibility."

Hydrogen has drawbacks as an energy source: It is less energy rich than natural gas and occupies large amounts of space. But experts think it could replace fossil fuels in long-haul transport and heavy industries such as steelmaking—if it could be sourced in an environmentally friendly way. Most hydrogen today is manufactured by combining steam and methane in factories that emit carbon dioxide (CO_2) and add to global warming. Governments are supporting efforts to make hydrogen cleanly, either by capturing the emitted CO_2 and storing it underground (blue hydrogen) or by using renewable electricity to split water and harvesting the resulting hydrogen (green hydrogen). Later this year, DOE plans to select up to 10 blue and green hydrogen "hubs" as part of an $8 billion program.



The well, in Nebraska, reached 3.4 kilometers down.

But "geological" or "natural" hydrogen could be cheaper—and just as clean. The ARPA-E money is a good start in a field with many unknowns, says Geoff Ellis, a geochemist at the U.S. Geological Survey (USGS) who leads a small, internally funded geological hydrogen program. "Right now, we're starting from, essentially, nearly zero. So this is really important as a way to jump-start research."

For decades, few geologists believed Earth held significant hydrogen deposits, because the gas is so readily eaten up by microbes or chemically altered into other forms. But prospectors are now fanning out across the globe, spurred by the discovery of a massive hydrogen field underneath a village in Mali and records suggesting puzzling surges of nearly pure hydrogen in old boreholes (*Science*, 17 February, p. 630). Whereas oil and gas companies tend to tap relatively youthful basins of sedimentary rock, hydrogen hunters are probing the crystalline, an-

cient hearts of continents for the iron-rich rocks thought to fuel hydrogen production.

The grant program will not support the hunt for existing deposits, because that is better left to USGS and industry, says ARPA-E Program Director Doug Wicks. Instead, it will focus on ways to artificially stimulate one of the main hydrogen producing reactions, called serpentinization, which occurs when water encounters iron-rich rocks at high temperatures and pressures. The reactions transform minerals such as olivine into serpentine, releasing hydrogen in the process. “How can we make the hydrogen come out faster?” Wicks asks.

Pumping water or catalysts to source rocks could help. One natural catalyst is magnetite, produced when olivine is serpentinized. The magnetite can catalyze subsequent hydrogen-producing reactions with the olivine, explains Alexis Templeton, a geochemist at the University of Colorado Boulder who plans to apply to the program. Some deep-living microbes can also catalyze hydrogen production, she says, and could be enlisted to speed it up. Much of the ARPA-E money is expected to go to modeling and lab-based research aimed at boosting production. But Templeton is glad to see the program will also fund research into monitoring the reactions and evaluating the risks of injecting substances into the crust. “What are we doing and how do we monitor and regulate that?”

Other backers are joining the hunt. In July, the natural hydrogen startup Koloma came out of stealth mode with $91 million in venture capital, including investments from Bill Gates’s Breakthrough Energy Ventures. Templeton just won a grant from the Grantham Foundation for the Protection of the Environment to explore ways to stimulate hydrogen production at lower temperatures and pressures. That could make it possible to generate the gas from the iron-rich deposits found near the surface in places such as Oman.

Ellis says major oil companies are now tentatively moving into the field after overlooking it for years. “We may be reaching a tipping point,” he says. This month, USGS and the Colorado School of Mines are launching a research consortium worth several million dollars over 5 years that includes support from companies including BP and Chevron.

Wicks says the investments by ARPA-E and USGS are “adding a level of credibility” to the fast-growing field. At first, he says, “I was a total skeptic.” But that changed after he and a colleague worked on hydrogen and dove into the literature. Now, he says, “I have a firm belief that it’s going to be a significant energy source in the future.” ■

PLANETARY SCIENCE

Did interstellar debris fall to the sea floor? Experts are doubtful

Controversial astrophysicist says metallic spheres hail from outside the Solar System, but others say it’s “nonsense”

By **Daniel Clery** *and* **Paul Voosen**

In 2014, a rock from space blazed through the atmosphere and exploded off the coast of Papua New Guinea with such ferociousness that some researchers believe the object came from beyond the Solar System. Now, a team of researchers says it has recovered remnants of the meteor from the floor of the Pacific Ocean and claims that a preliminary analysis of their unusual composition points to an origin around another star.

Only in the past few years have astronomers realized that interstellar objects sometimes whizz through the Solar System and might even hit Earth. Finding a lump of rock from another planetary system would be an unbelievable stroke of scientific fortune, one that could shed light on the formation of alien planets and stars.

> **“If you don’t allow for surprises, you won’t learn something new.”**
> **Avi Loeb**, Harvard University

Avi Loeb, the controversial theoretical physicist at Harvard University who led the team, believes that’s what his high-risk ocean mission has achieved. “If you don’t allow for surprises, you won’t learn something new,” he says. On 29 August, the team released a preprint describing the claims, which it has submitted to the journal *Ocean Science*.

But others are dismissive of the preprint, which has not been peer reviewed. Although the geochemical analysis of the debris is solid, the conclusions that Loeb and his colleagues hang on it are “nonsense,” says Martin Schiller, a cosmochemist at the University of Copenhagen. “I’m surprised anyone would take it seriously.” Larry Nittler, a cosmochemist at Arizona State University (ASU), calls the claims “very weak sauce.”

Loeb has gained notoriety in recent years for his view that ‘Oumuamua, a body that passed through the Solar System in 2017 on a path that suggested an interstellar origin, might be an alien spacecraft. In 2021, he launched the Galileo Project, a privately funded effort to use scientific methods to search for evidence of alien technology on or near Earth.

The origin of his latest project lies in his discovery that the U.S. Department of Defense’s Space Command keeps a catalog of meteors detected by its surveillance satellites as they explode in Earth’s atmosphere. Loeb and his student Amir Siraj studied the 272 events in the catalog to gauge whether any of the meteors were on a trajectory from outside the Solar System. One seemed to fit the bill. It had plummeted into the atmosphere just north of Manus Island in Papua New Guinea on 8 January 2014 at a speed of 45 kilometers per second—faster than any object orbiting the Sun—before shattering in three explosions.

In 2022, Space Command reported in a letter to NASA that it was 99.999% certain the meteor was from beyond the Solar System. It also released details about the three explosions showing the meteor deeply penetrated the atmosphere before breaking up, which Loeb says indicates it was even tougher than the iron objects from the Solar System that routinely strike Earth. “It was an outlier in material strength,” Loeb says. Space Command located the meteor’s track to an 11-kilometer-wide square of ocean. With the help of a seismometer on Manus, which recorded the explosions, Loeb and Siraj narrowed down its location to a 1-kilometer-wide strip.

Loeb acquired $1.5 million from cryptocurrency entrepreneur Charles Hoskinson to go look for the debris. In June, the *M/V Silver Star* began to trawl the ocean bottom off Manus. Over 2 weeks, Loeb and colleagues dragged a sled covered with neodymium magnets along the sea floor, 2 kilometers down, hoping the magnets would pick up metallic meteor fragments. After each 8-hour run, the magnets were scraped and vacuumed clean. Loeb says they were lucky to have found anything. “There could have been so many failure points.”

On the ship and back at Harvard, the team picked over the debris with tweezers and found, amid volcanic ash, nearly 700 “spherules,” tiny metallic pellets 1 millimeter or less across. Droplets of

Check for updates

cient hearts of continents for the iron-rich rocks thought to fuel hydrogen production.

The grant program will not support the hunt for existing deposits, because that is better left to USGS and industry, says ARPA-E Program Director Doug Wicks. Instead, it will focus on ways to artificially stimulate one of the main hydrogen producing reactions, called serpentinization, which occurs when water encounters iron-rich rocks at high temperatures and pressures. The reactions transform minerals such as olivine into serpentine, releasing hydrogen in the process. "How can we make the hydrogen come out faster?" Wicks asks.

Pumping water or catalysts to source rocks could help. One natural catalyst is magnetite, produced when olivine is serpentinized. The magnetite can catalyze subsequent hydrogen-producing reactions with the olivine, explains Alexis Templeton, a geochemist at the University of Colorado Boulder who plans to apply to the program. Some deep-living microbes can also catalyze hydrogen production, she says, and could be enlisted to speed it up. Much of the ARPA-E money is expected to go to modeling and lab-based research aimed at boosting production. But Templeton is glad to see the program will also fund research into monitoring the reactions and evaluating the risks of injecting substances into the crust. "What are we doing and how do we monitor and regulate that?"

Other backers are joining the hunt. In July, the natural hydrogen startup Koloma came out of stealth mode with $91 million in venture capital, including investments from Bill Gates's Breakthrough Energy Ventures. Templeton just won a grant from the Grantham Foundation for the Protection of the Environment to explore ways to stimulate hydrogen production at lower temperatures and pressures. That could make it possible to generate the gas from the iron-rich deposits found near the surface in places such as Oman.

Ellis says major oil companies are now tentatively moving into the field after overlooking it for years. "We may be reaching a tipping point," he says. This month, USGS and the Colorado School of Mines are launching a research consortium worth several million dollars over 5 years that includes support from companies including BP and Chevron.

Wicks says the investments by ARPA-E and USGS are "adding a level of credibility" to the fast-growing field. At first, he says, "I was a total skeptic." But that changed after he and a colleague worked on hydrogen and dove into the literature. Now, he says, "I have a firm belief that it's going to be a significant energy source in the future." ■

PLANETARY SCIENCE

Did interstellar debris fall to the sea floor? Experts are doubtful

Controversial astrophysicist says metallic spheres hail from outside the Solar System, but others say it's "nonsense"

By **Daniel Clery** *and* **Paul Voosen**

In 2014, a rock from space blazed through the atmosphere and exploded off the coast of Papua New Guinea with such ferociousness that some researchers believe the object came from beyond the Solar System. Now, a team of researchers says it has recovered remnants of the meteor from the floor of the Pacific Ocean and claims that a preliminary analysis of their unusual composition points to an origin around another star.

Only in the past few years have astronomers realized that interstellar objects sometimes whizz through the Solar System and might even hit Earth. Finding a lump of rock from another planetary system would be an unbelievable stroke of scientific fortune, one that could shed light on the formation of alien planets and stars.

> **"If you don't allow for surprises, you won't learn something new."**
> **Avi Loeb**, Harvard University

Avi Loeb, the controversial theoretical physicist at Harvard University who led the team, believes that's what his high-risk ocean mission has achieved. "If you don't allow for surprises, you won't learn something new," he says. On 29 August, the team released a preprint describing the claims, which it has submitted to the journal *Ocean Science*.

But others are dismissive of the preprint, which has not been peer reviewed. Although the geochemical analysis of the debris is solid, the conclusions that Loeb and his colleagues hang on it are "nonsense," says Martin Schiller, a cosmochemist at the University of Copenhagen. "I'm surprised anyone would take it seriously." Larry Nittler, a cosmochemist at Arizona State University (ASU), calls the claims "very weak sauce."

Loeb has gained notoriety in recent years for his view that 'Oumuamua, a body that passed through the Solar System in 2017 on a path that suggested an interstellar origin, might be an alien spacecraft. In 2021, he launched the Galileo Project, a privately funded effort to use scientific methods to search for evidence of alien technology on or near Earth.

The origin of his latest project lies in his discovery that the U.S. Department of Defense's Space Command keeps a catalog of meteors detected by its surveillance satellites as they explode in Earth's atmosphere. Loeb and his student Amir Siraj studied the 272 events in the catalog to gauge whether any of the meteors were on a trajectory from outside the Solar System. One seemed to fit the bill. It had plummeted into the atmosphere just north of Manus Island in Papua New Guinea on 8 January 2014 at a speed of 45 kilometers per second—faster than any object orbiting the Sun—before shattering in three explosions.

In 2022, Space Command reported in a letter to NASA that it was 99.999% certain the meteor was from beyond the Solar System. It also released details about the three explosions showing the meteor deeply penetrated the atmosphere before breaking up, which Loeb says indicates it was even tougher than the iron objects from the Solar System that routinely strike Earth. "It was an outlier in material strength," Loeb says. Space Command located the meteor's track to an 11-kilometer-wide square of ocean. With the help of a seismometer on Manus, which recorded the explosions, Loeb and Siraj narrowed down its location to a 1-kilometer-wide strip.

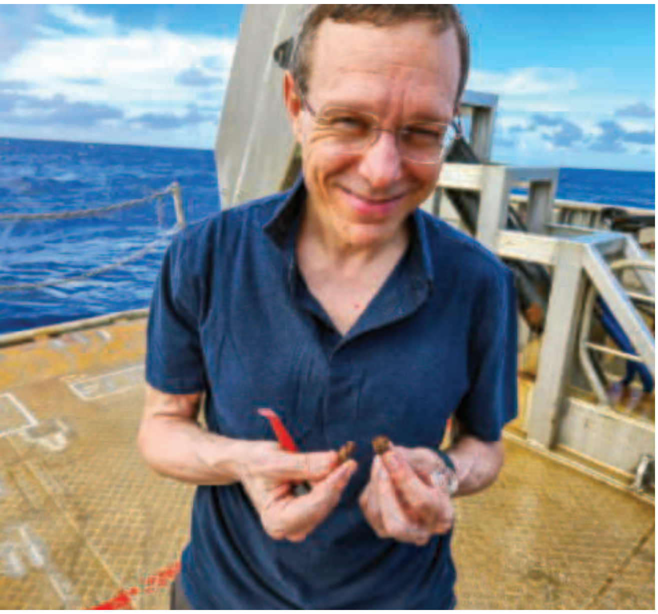
Loeb acquired $1.5 million from cryptocurrency entrepreneur Charles Hoskinson to go look for the debris. In June, the *M/V Silver Star* began to trawl the ocean bottom off Manus. Over 2 weeks, Loeb and colleagues dragged a sled covered with neodymium magnets along the sea floor, 2 kilometers down, hoping the magnets would pick up metallic meteor fragments. After each 8-hour run, the magnets were scraped and vacuumed clean. Loeb says they were lucky to have found anything. "There could have been so many failure points."

On the ship and back at Harvard, the team picked over the debris with tweezers and found, amid volcanic ash, nearly 700 "spherules," tiny metallic pellets 1 millimeter or less across. Droplets of

once-molten material, spherules routinely form in explosions of ordinary meteors, and also in volcanic eruptions.

Loeb delivered spherules to several labs for compositional analysis, including one run by Harvard geochemist Stein Jacobsen. Jacobsen found that the ratios of iron isotopes in the spherules largely matched those of the Sun—a mark against an interstellar origin. However, five of the spherules were unusually enriched in beryllium and lanthanum, and, to a lesser degree, uranium. This "BeLaU" fingerprint hasn't been seen in records of known meteoritic spherules, Jacobsen says, although it does resemble some lunar samples. "The fact is we found something unusual that has not been seen."

Beryllium is rare in the universe, and most is produced when cosmic rays collide with large atoms, chipping off atomic fragments such as beryllium. On a long interstellar journey, an object would have more exposure to cosmic rays than any rock from the Solar System, and more time to produce beryllium, Loeb says. "Is beryllium a flag for interstellar travel?"



Avi Loeb claims to have found remnants of an interstellar meteor on the seabed of the Pacific Ocean, but experts are doubtful.

The three enriched elements also tend to bind with iron. The authors suggest they may have crystallized from a magma ocean that covered the surface of a primitive celestial body with an iron core.

Jacobsen himself acknowledges the pattern does not on its own indicate that the spherules came from outside the Solar System. Planetary embryos with an iron core could have been numerous in the early days of the Solar System. Many known meteorites are thought to have such an origin, even if their detailed composition isn't known. "Occam's razor," Schiller says. "There is no evidence it comes from outside the Solar System."

Indeed, the team needs to do more to prove the spherules came from space and are not volcanic, Nittler says. So far, only one of the anomalous spherules has the isotopic pattern in its iron expected if it was heated by a fiery passage through the atmosphere. The paper also did not consider how volcanic eruptions can interact with other rocks to create strange combinations of elements, says Frédéric Moynier, a cosmochemist at the Paris Institute of Planetary Physics. Given the claims, he adds, "I don't think it would pass any thorough review process."

Finally, the central pillar of Loeb's argument that the meteor was interstellar—its high speed—does not seem to be as certain as Space Command's 99.999% estimate. A new study out this month in *The Astrophysical Journal* examined 17 known fireballs captured by both classified U.S. sensors and independent observations. The study showed that the government sensors often overestimated speed, with the errors getting worse the faster things got. "A third of the time, the numbers are just way off," says Steve Desch, an ASU astrophysicist.

Desch says a meteor hitting the atmosphere at 45 kilometers per second would probably be completely vaporized, leaving barely any solid debris at all. He also says the link between the spherules and the 2014 fireball is tenuous. Even if the spherules hit the ocean where Loeb says, they would have likely drifted by tens of kilometers in ocean currents before settling on the seabed, Desch argues. Loeb's team collected too few control samples from elsewhere on the sea floor to be sure its finds were unusual, Desch says. "They knew what they were looking for and that makes it prone to confirmation bias."

Jacobsen says the spherules could yield more clues. He wants to look for other isotopic variations in trace elements such as neodymium, which could indicate formation around another star. But given the spherules' small size—one weighed only 27 micrograms—teasing out that signal could be a challenge.

Jacobsen says he typically does much more work before submitting a paper. "It's not my style—normally I work on something a bit longer." But Loeb wanted to do something quickly, "so that's what we did," he says. ■

U.S. RESEARCH POLICY

Budget fight raises fears of spending cuts, research limits

Lawmakers return from recess to full plate of issues

By ***Science*** **News Staff**

A flat budget isn't something that scientists typically welcome. But for the U.S. research community, that's the best-case scenario in the short run as Congress returns this month from its summer recess and tries to agree on spending levels for the 2024 fiscal year that begins on 1 October. The government could shut down if Republicans and Democrats don't declare a truce.

The next federal budget is one of several topics affecting science that await action by this divided Congress, where Democrats hold a narrow edge in the Senate and Republicans have an equally small majority in the House of Representatives. Legislators are also debating proposals for greater oversight of research collaborations with China and restrictions on research that alters the characteristics of pathogens, as well as a bill setting new funding targets for agricultural research. Biomedical scientists are also hoping the Senate will confirm Monica Bertagnolli, nominated in May, as the new head of the National Institutes of Health (NIH).

The 2024 budget is the most contentious issue facing lawmakers. With no prospect of passing all 12 bills that set new spending levels by 1 October, Democratic leaders in both houses and many Republicans are calling for a continuing resolution (CR) that would freeze budgets at this year's levels for a few months while negotiations continue. But some three dozen staunch conservatives in the House have pledged to oppose a CR. They want Congress to roll back spending to 2022 levels, which could cut some research programs by 20% or more.

Democrats vow not to let that happen. If Congress can't pass a CR by 30 September, however, most government activities would grind to a halt.

Such temporary shutdowns are not uncommon. There have been four since 1995, including for 35 days in 2018–19. Once those



Congress will be wrestling with fiercely partisan legislation this month that affects research.

shutdowns ended, Congress ultimately approved catchall spending bills that usually included small increases for research.

Achieving that benign outcome will be harder this year. In May, a deal to avoid defaulting on the $31 trillion federal debt included a provision that would reduce federal spending by 1% from this year's levels if Congress failed to approve a 2024 budget for all agencies by 30 April 2024.

Given that pressure to hold down government spending, science advocates don't expect any agency to get more than a minimal increase at best when Congress finally passes a 2024 budget. For example, NIH would get a 2% boost, to $47.8 billion, if Congress eventually adopts a Senate proposal. But if a House spending panel has its way, NIH would get a 6% cut. At the same time, committees in both houses have proposed giving the National Science Foundation some $300 million less than the $9.87 billion it received this year.

Here are some other important issues pending before Congress.

FARM BILL

A 2018 law, which expires on 31 September, set funding priorities and policies at the U.S. Department of Agriculture. In addition to supplying food aid to low-income residents and crop insurance to farmers, the law provided $694 million over 5 years for research on plant genetics, crop diseases, and soil health.

Research advocates hope the new bill will include $5 billion to repair and upgrade laboratories and other research infrastructure. "It really has to be a total overhaul of the system," says Doug Steele of the Association of Public & Land-grant Universities. "We're almost at the breaking point." But progress on the bill has been slow, in part because fiscal conservatives don't want to increase the $725 billion in mandatory spending in the expiring legislation.

NIH DIRECTOR

The nomination of Bertagnolli to replace Francis Collins, who stepped down in December 2021, got a boost last week after one of two Democratic members on the Senate health committee that oversees NIH dropped their opposition. Bertagnolli, a surgical oncologist who now leads NIH's National Cancer Institute, promised Senator Elizabeth Warren (D–MA) not to take a job with a major drug company for 4 years after she leaves government. But Senator Bernie Sanders (I–VT), the panel's chair, has said he won't schedule a vote on the nomination until President Joe Biden's administration takes more steps to lower drug prices.

"I'm hopeful that [Sanders] can be convinced an agency the size and importance of NIH needs to have a permanent head, someone who can do more than simply hold down the fort," says Carrie Wolinetz, a lobbyist for Lewis-Burke Associates and a former senior NIH official. Sanders's office did not respond to a request for comment.

GAIN-OF-FUNCTION (GOF) RESEARCH

Several House bills would ban federal funding for GOF studies that alter the characteristics of pathogens. The measures, some of which focus on lab research in China and other "adversary" countries, are backed by conservatives who have suggested that U.S.-funded studies on bat coronaviruses in Wuhan, China, created the SARS-CoV-2 pandemic virus. But microbiologists fear vaguely worded provisions could shut down studies in the United States and abroad aimed at preparing the world for the next pandemic.

"Cutting off research in the U.S. could endanger us by encouraging the work to be done in places that have less stringent standards," says Allen Segal, chief advocacy officer for the American Society for Microbiology. Some bills would also ban funding for the EcoHealth Alliance, a U.S. nonprofit that partnered with virologists in Wuhan.

RESEARCH SECURITY

Both the House and Senate have added provisions to their versions of an annual bill that sets policy for the Department of Defense (DOD) that are aimed at preventing adversaries from using U.S.-funded research to undermine national security. Many specifically target China. One House proposal, for instance, would require the Department of the Treasury to produce a report "on the extent to which China has benefitted from United States taxpayer-funded research."

Research advocates are most concerned about a House provision that would require anyone working on a DOD-funded grant to disclose detailed personal information and work histories, which would be posted on a public website. They argue the requirement would place a heavy additional burden on institutions and be counterproductive by providing adversaries with a guide to DOD grantees and their research.

OPEN ACCESS

Advocates for making scientific articles free to read are hoping Congress will reject a House spending panel's proposal that would block a White House directive supporting open access. The proposal would prevent using federal funds to implement or enforce White House guidance issued last year requiring research articles produced with federal funding to be placed in a public repository as soon as they are published. Previous White House policy allowed for a 12-month embargo, a delay that benefits publishers who rely on revenue from paywalled content. ■

With reporting by Jeffrey Brainard, Jocelyn Kaiser, Jeffrey Mervis, and Erik Stokstad.

PHOTO: NATHAN HOWARD/BLOOMBERG VIA GETTY IMAGES



This moth, *Neopalpa donaldtrumpi*, was named after the former U.S. president.

TAXONOMY

Should beetles be named after Adolf Hitler?

Zoologists debate whether—and how—to change scientific names now deemed offensive

By **Rodrigo Pérez Ortega**

In 1934, a German paleontologist named a giant flying insect from the Carboniferous period *Rochlingia hitleri*, after Adolf Hitler, who had just taken power in Germany, and Hermann Röchling, an antisemitic steel manufacturer and member of the Nazi Party. Three years later, an Austrian amateur entomologist named a brown, eyeless beetle from Slovenian caves *Anophthalmus hitleri* because he admired Hitler. In recent years, neo-Nazis have reportedly paid thousands for specimens, pushing the beetle toward extinction.

Some researchers have argued for years that *A. hitleri* and other species names, including the many that honor racists and colonizers, are offensive and should be changed. A few societies have taken steps toward doing so. But not the International Commission on Zoological Nomenclature (ICZN), and its stance has ignited fierce debate.

In January, the commission, which arbitrates on the correct use of scientific names of animals, announced in the *Zoological Journal of the Linnean Society* (*ZJLS*) that it will not consider changing animal names many researchers consider offensive. "If these names are not stable, you can create a massive confusion," explains ICZN Commissioner Luis Ceríaco, a biologist at the University of Porto. But on 23 August, a series of editorials in the same journal pushed back, saying the decision was made without feedback from the community and wrongly prioritized tradition over ethics. It's a matter of "eliminating the commemoration of people who caused untold human misery," says one author, botanist Estrela Figueiredo of Nelson Mandela University's Ria Olivier Herbarium. "In which other spheres of human endeavor is anything still named [after] Hitler? … The codes must change and adapt, like the rest of society."

Names identified as problematic include *Hypopta mussolinii*, a butterfly discovered in Libya and named after Benito Mussolini, the fascist Italian leader who invaded the country. And sometimes, organisms are named apparently with intent to mock: In 2017, researchers named a moth with pale blond head scales and small genitalia *Neopalpa donaldtrumpi*.

The ICZN commissioners, who are in charge of the International Code of Zoological Nomenclature, argued in January that renaming animals because of cultural offense would disrupt the code's chief goal: stability. Scientists would then use more than one name to refer to the same species.



Researchers argue that *Anophthalmus hitleri* needs a new scientific name.

But the editorial authors disagree. "We cannot prioritize stability over social justice," counters Marcos Raposo, an ornithologist and taxonomist at the National Museum of the Federal University of Rio de Janeiro, author of another editorial.

Debate over scientific names erupted earlier among botanists, who have traded editorials and letters over the past 2 years. Some pushed to restore Indigenous species names and change offensive ones, such as *Hibbertia*, a large genus of Australian guinea flowers that honors English antiabolitionist and plantation owner George Hibbert. A few botanists proposed getting rid of eponyms honoring humans altogether. Other researchers pushed back, saying the nomenclature shouldn't be biased by social concerns.

The discussion got heated, but "the botanical community is trying really hard to be grown up about this," says botanist Sandra Knapp of the Natural History Museum (NHM) in London. She is also president of the nomenclature section of the 2024 International Botanical Congress in Madrid, and notes that the naming code for plants, animals, and fungi allows voting by members: At the meeting in July, botanists will vote on a proposal to allow existing culturally offensive names to be changed and to form a committee to do so.

As for animals, some scientific societies have already moved to change common names that give offense. In 2022, for in-

stance, the Entomological Society of America adopted spongy moth for the invasive moth *Lymantria dispar*, getting rid of gypsy moth. But the ICZN commissioners argue scientific names are a bigger challenge.

They estimated in the January statement that about 20% of the more than 1.5 million animal names are eponyms and about 10% are toponyms, referring to a place. That suggests that several hundred thousand scientific names could be challenged on ethical grounds. Adjudicating the challenges and deciding on new names would be a nearly impossible task, Ceríaco says.

In a third editorial, Christopher Bae, a paleoanthropologist at the University of Hawaii at Manoa, and colleagues suggested creating a special ethics committee to revise names, thus spreading the workload. But that's beyond the commission's scope, Ceríaco says. "None of us were elected to be commissioners [because of our] expertise on ethics," he says. "We are experts on nomenclature and taxonomy."

Figueiredo, Raposo, and others argue that ICZN should invite outside views. "They should ask our opinion," Raposo says, and give priority to members of historically oppressed groups. He and other critics note that most ICZN commissioners are white and based in the Global North. None is from Africa, where about 1500 vertebrate species have eponyms, many reflecting the continent's history of imperialism.

Keeping names with racist slurs denies dignity to marginalized communities, agrees paleontologist Anjali Goswami of NHM. She is also the president of the Linnean Society, which publishes *ZJLS*. "There are ways that we can at least start to make some progress on the worst offensive [names] and then move towards a better practice."

For example, for plants, Figueiredo and Gideon Smith, also at Nelson Mandela University, say some offensive names could be eliminated in a single stroke, by removing the "c" in about 300 botanical names from Africa with the roots *caf[e]r-* and *caff[e]r-*. The one-letter change would transform an allusion to a now illegal, Apartheid-era slur used to discriminate against Black people in South Africa to *afer-*, implying a species origin in Africa. The change would cause little disturbance to the nomenclature, the researchers say. "The ICZN needs to be receptive to proposals for change," Figueiredo says.

ICZN acknowledges the troubled past of some names, and will take note of the concerns now published, Ceríaco says. "It's good for us to now read and absorb what the community is discussing." As for *A. hitleri*, given that the beetle's current name threatens its survival, Ceríaco says, ICZN might consider giving it a new one. ■

U.S. SCIENCE POLICY

Researchers applaud HHS push to ease cannabis restrictions

Proposed rescheduling would make it easier for scientists to acquire and study the drug

By **Phie Jacobs**

Federal health officials are urging the U.S. Drug Enforcement Administration (DEA) to loosen its restrictions on cannabis—a move that could make it easier for researchers to study the drug's potential medical benefits and harms. Following a review initiated by the White House in 2022, the U.S. Department of Human Health and Services (HHS) last week recommended that DEA reclassify cannabis from its Schedule I category, which includes drugs considered to have a high potential for abuse and no accepted therapeutic value, such as heroin and LSD, to the lower risk Schedule III. If implemented, the policy change could relax lengthy licensing and handling procedures that scientists say have hampered research.

"This is a really unprecedented situation," says biopsychologist Ziva Cooper, director of the Center for Cannabis and Cannabinoids at the University of California (UC), Los Angeles. Nearly half of U.S. states have legalized recreational cannabis, and medicinal use is permitted in many more. Now, after decades of resisting calls from scientists and activists to reschedule the drug, DEA faces new pressure to do so.

The HHS recommendation, sent to DEA in a letter first reported by Bloomberg News, would place cannabis in the same category as ketamine and anabolic steroids, which are used in health care settings and can be obtained with a prescription. (A DEA spokesperson told Bloomberg the agency had received the letter and will now initiate its own review process.)

Cannabis has shown promise in relieving chronic pain and is being explored as a possible treatment for cancer, post-traumatic stress disorder, and other conditions. "From a medical perspective, it fits better" in Schedule III, says psychiatrist Igor Grant, director of the Center for Medicinal Cannabis Research at UC San Diego.

Scientists who study cannabis have long chafed under DEA policies. The Schedule I classification "makes everything more challenging," says neuroscientist Staci Gruber, who runs trials involving cannabis at McLean Hospital in Massachusetts. A 2022 law increased access to cannabis for medical research (*Science*, 9 December 2022, p. 1035), but scientists must still apply for a DEA license—a process that requires months of paperwork and must be repeated for each new study. Rescheduling would streamline the process. For example, a research manual published by DEA indicates that scientists working with Schedule III substances aren't required to submit their study protocols to the agency in advance.

Researchers also expect an easing of security rules for storage and handling, which currently require them to use high-tech lock boxes and expensive security cameras.

The proposed change could increase the supply of cannabis for research. Currently, DEA only permits a few universities and companies to produce the plant. Obtaining a permit means investing in a complicated security system and hiring trained guards, says George Hodgin, CEO of the Biopharmaceutical Research Company, which provides cannabis to researchers. Rescheduling wouldn't produce "an overnight sea change," he says, but it could loosen the requirements enough to bring in more suppliers.

The DEA manual also notes that researchers working with substances in Schedules II through V can apply to produce small quantities of product rather than purchasing it from an approved facility.

Whether DEA will approve HHS's recommendation isn't clear. But it often defers to HHS on scientific and medical matters, notes attorney Larry Houck, who spent 15 years as a DEA diversion investigator.

In the meantime, researchers have a slew of questions about cannabis to address. "We're still a long way from having enough empirically sound data" to support its therapeutic benefits, Gruber says. Its growing availability has made research more urgent. According to the U.S. Centers for Disease Control and Prevention, one in five Americans used the drug in 2019, and new cannabis products, some with a high tetrahydrocannabinol content, are flooding the market. "We have a responsibility to try to understand the ways in which [people] may use it more responsibly," Gruber says. ■

stance, the Entomological Society of America adopted spongy moth for the invasive moth *Lymantria dispar*, getting rid of gypsy moth. But the ICZN commissioners argue scientific names are a bigger challenge.

They estimated in the January statement that about 20% of the more than 1.5 million animal names are eponyms and about 10% are toponyms, referring to a place. That suggests that several hundred thousand scientific names could be challenged on ethical grounds. Adjudicating the challenges and deciding on new names would be a nearly impossible task, Ceríaco says.

In a third editorial, Christopher Bae, a paleoanthropologist at the University of Hawaii at Manoa, and colleagues suggested creating a special ethics committee to revise names, thus spreading the workload. But that's beyond the commission's scope, Ceríaco says. "None of us were elected to be commissioners [because of our] expertise on ethics," he says. "We are experts on nomenclature and taxonomy."

Figueiredo, Raposo, and others argue that ICZN should invite outside views. "They should ask our opinion," Raposo says, and give priority to members of historically oppressed groups. He and other critics note that most ICZN commissioners are white and based in the Global North. None is from Africa, where about 1500 vertebrate species have eponyms, many reflecting the continent's history of imperialism.

Keeping names with racist slurs denies dignity to marginalized communities, agrees paleontologist Anjali Goswami of NHM. She is also the president of the Linnean Society, which publishes *ZJLS*. "There are ways that we can at least start to make some progress on the worst offensive [names] and then move towards a better practice."

For example, for plants, Figueiredo and Gideon Smith, also at Nelson Mandela University, say some offensive names could be eliminated in a single stroke, by removing the "c" in about 300 botanical names from Africa with the roots *caf[e]r-* and *caff[e]r-*. The one-letter change would transform an allusion to a now illegal, Apartheid-era slur used to discriminate against Black people in South Africa to *afer-*, implying a species origin in Africa. The change would cause little disturbance to the nomenclature, the researchers say. "The ICZN needs to be receptive to proposals for change," Figueiredo says.

ICZN acknowledges the troubled past of some names, and will take note of the concerns now published, Ceríaco says. "It's good for us to now read and absorb what the community is discussing." As for *A. hitleri*, given that the beetle's current name threatens its survival, Ceríaco says, ICZN might consider giving it a new one. ■

U.S. SCIENCE POLICY

Researchers applaud HHS push to ease cannabis restrictions

Proposed rescheduling would make it easier for scientists to acquire and study the drug

By **Phie Jacobs**

Federal health officials are urging the U.S. Drug Enforcement Administration (DEA) to loosen its restrictions on cannabis—a move that could make it easier for researchers to study the drug's potential medical benefits and harms. Following a review initiated by the White House in 2022, the U.S. Department of Human Health and Services (HHS) last week recommended that DEA reclassify cannabis from its Schedule I category, which includes drugs considered to have a high potential for abuse and no accepted therapeutic value, such as heroin and LSD, to the lower risk Schedule III. If implemented, the policy change could relax lengthy licensing and handling procedures that scientists say have hampered research.

"This is a really unprecedented situation," says biopsychologist Ziva Cooper, director of the Center for Cannabis and Cannabinoids at the University of California (UC), Los Angeles. Nearly half of U.S. states have legalized recreational cannabis, and medicinal use is permitted in many more. Now, after decades of resisting calls from scientists and activists to reschedule the drug, DEA faces new pressure to do so.

The HHS recommendation, sent to DEA in a letter first reported by Bloomberg News, would place cannabis in the same category as ketamine and anabolic steroids, which are used in health care settings and can be obtained with a prescription. (A DEA spokesperson told Bloomberg the agency had received the letter and will now initiate its own review process.)

Cannabis has shown promise in relieving chronic pain and is being explored as a possible treatment for cancer, posttraumatic stress disorder, and other conditions. "From a medical perspective, it fits better" in Schedule III, says psychiatrist Igor Grant, director of the Center for Medicinal Cannabis Research at UC San Diego.

Scientists who study cannabis have long chafed under DEA policies. The Schedule I classification "makes everything more challenging," says neuroscientist Staci Gruber, who runs trials involving cannabis at McLean Hospital in Massachusetts. A 2022 law increased access to cannabis for medical research (*Science*, 9 December 2022, p. 1035), but scientists must still apply for a DEA license—a process that requires months of paperwork and must be repeated for each new study. Rescheduling would streamline the process. For example, a research manual published by DEA indicates that scientists working with Schedule III substances aren't required to submit their study protocols to the agency in advance.

Researchers also expect an easing of security rules for storage and handling, which currently require them to use high-tech lock boxes and expensive security cameras.

The proposed change could increase the supply of cannabis for research. Currently, DEA only permits a few universities and companies to produce the plant. Obtaining a permit means investing in a complicated security system and hiring trained guards, says George Hodgin, CEO of the Biopharmaceutical Research Company, which provides cannabis to researchers. Rescheduling wouldn't produce "an overnight sea change," he says, but it could loosen the requirements enough to bring in more suppliers.

The DEA manual also notes that researchers working with substances in Schedules II through V can apply to produce small quantities of product rather than purchasing it from an approved facility.

Whether DEA will approve HHS's recommendation isn't clear. But it often defers to HHS on scientific and medical matters, notes attorney Larry Houck, who spent 15 years as a DEA diversion investigator.

In the meantime, researchers have a slew of questions about cannabis to address. "We're still a long way from having enough empirically sound data" to support its therapeutic benefits, Gruber says. Its growing availability has made research more urgent. According to the U.S. Centers for Disease Control and Prevention, one in five Americans used the drug in 2019, and new cannabis products, some with a high tetrahydrocannabinol content, are flooding the market. "We have a responsibility to try to understand the ways in which [people] may use it more responsibly," Gruber says. ■



OFF THE GRID

Computer models that forecast overloaded power lines are holding back U.S. solar and wind energy projects

By **Dan Charles**, *in Little Rock, Arkansas*

One morning earlier this year, all seemed calm in the dimly lit, bunkerlike control room of the Southwest Power Pool (SPP) here, which manages an electricity grid that stretches across the high plains from North Dakota to New Mexico. Wall-size displays showed electricity flowing smoothly through SPP's 100,000 kilometers of transmission lines, with 80% of the power coming from plants fueled by coal and natural gas.

The control room's tranquility, however, belied an emerging, high-stakes battle to redraw this map. SPP and other U.S. grid operators are facing an unprecedented tsunami of requests from energy firms to connect thousands of proposed wind, solar, and power storage projects to their transmission lines. The projects are essential to meeting the U.S. goal of eliminating all planet-warming carbon emissions from the nation's electricity supply by 2035, analysts say. Together, they could generate almost 2000 gigawatts of electricity—exceeding the total capacity of the country's existing power plants.

Most of these projects, however, have been stuck in limbo for years, waiting in what energy insiders call the "interconnection queue." One contributor to the bottleneck: mathematical simulations that SPP and other operators use to predict how electricity from those new power generators will affect the grid's stability and reliability.

These simulations do not address the much-debated complications of relying for a steady stream of electricity on wind and solar power sources, which are subject to the whims of weather and deliver power intermittently. That remains a concern, but it doesn't prevent new projects from coming online. Interconnection studies are focused on a different question: Will the projects generate too much power, and in the wrong places, testing the limits of what power lines can handle before they start to melt? To avoid such damaging congestion, grid operators require renewable power producers to pay up front for expensive transmission upgrades. But many can't afford those improvements and must abandon their plans.

"Interconnection is becoming one of the leading barriers to bringing projects online," says Joe Rand, a researcher at Lawrence Berkeley National Laboratory who tracks projects in the interconnection queue.



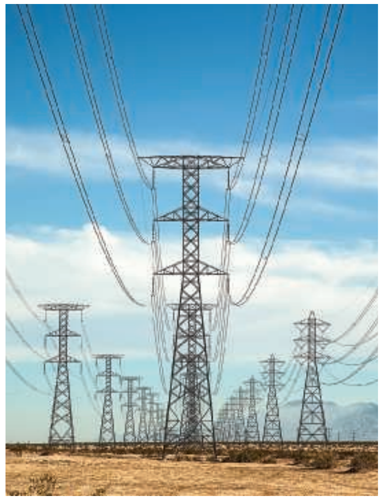
Staffers monitor 100,000 kilometers of power lines at a control room operated by the Southwest Power Pool.

Some researchers and renewable power advocates argue that the interconnection logjam is, in part, a product of flawed simulations based on assumptions that are too conservative and sometimes unreasonable. "These assumptions are so important, because they ultimately dictate" what it costs to build new generation, says Aaron Vander Vorst, head of growth strategy and transmission at Enel North America, a major developer of wind and solar projects. Yet they give rise to show-stopping scenarios of overloaded lines that don't reflect reality, he and others say.

Grid managers like SPP insist that interconnection simulations are essential to making sure new renewable power sources don't overwhelm the grid. Yet, Vander Vorst and others argue that they have become a kind of malfunctioning toll gate on the road toward clean energy. To fix that toll gate, some want to see the adoption of more realistic model assumptions and more flexible rules. And when new power lines really are needed, they say, the burden should not fall entirely on the new green power projects.

THE INTERCONNECTION roadblock is personal for farm families who live amid the windswept fields and grasslands of eastern New Mexico. "We have the best wind in the country, you know," says Eva Woods, a resident of Broadview, New Mexico.

A decade ago, Apex Clean Energy, a Virginia-based firm, drew up plans to harvest that power. Local farmers were ready to lease their land as sites for up to 135 wind turbines that together could generate 300 megawatts of electricity—enough to power 100,000 homes. Regular income from such leases "helps our farmers and ranchers so much," says Woods, a well-connected figure who's been publicly supporting the project.



Renewable energy projects often face long waits to connect to transmission lines.

In early 2017, Apex applied for permission to connect this potential project, called Grady Martin Wind, to a power line just outside the town of Clovis. The request went to SPP, which manages the region's grid.

Grady Martin was just one of dozens of wind and solar projects that asked SPP for interconnection around the same time. Like many grid operators, SPP struggled to keep up with the surge of applications, and it took years to launch the needed study. But in 2021, SPP's engineers were able to fire up what's called a power flow model, which calculates the flow of electricity through every power line and transformer on the grid. They ran a simulation in which 60 wind and solar farms—including Grady Martin—replaced an equivalent amount of power from existing power plants elsewhere. This amounted to a major rerouting of electricity; the proposed projects represented about 10,000 megawatts of generation, equivalent to about 20% of the peak demand for power in SPP's region on a hot summer day.

The model choked on the glut of new renewable energy. Its equations could not find a solution to meeting the region's power demands with this new fleet of generators, using existing power lines. With all the wind and solar plants generating maximum power, the system couldn't safely handle it all.

And the model hadn't even gotten to another essential part of the study, which involved simulating what would happen if a routine problem—such as the failure of a single power line—occurred.

To relieve the congestion, SPP's engineers had to add hefty new power lines to the simulation. Their most notable addition: what's known as a double-circuit 765kV line running more than 200 kilometers southeast from Lubbock, Texas. Only a handful of power lines this big exist in the United States, none of them west of the Mississippi. "A 765kV line is like a 10-lane highway through Los Angeles, right?" says David Kelley, SPP's vice president of engineering.

In the model, adding those power pipes relieved congestion and allowed all 60 wind and solar projects to connect to the grid. But there were practical hitches: Building that new 765kV line, for instance, would cost well over $1 billion, SPP estimated. And it assigned that cost to 14 proposed wind and solar farms, including Apex's Grady Martin. Proposed projects were also expected to pay for more routine fixes. In total, the projects were on the hook for $4.6 billion in transmission upgrades. Grady Martin's share was $272 million.

Faced with such costs, Apex withdrew its application. So did half of the 60-odd projects in the study group. The dropouts included 13 of the 14 projects that were supposed to pay for the 765kV line. When those projects evaporated, so did the need for the electricity superhighway. "That 765kV line is not getting built," Kelley says. Nor is Grady Martin, at least for now—and farm families are getting no checks.

THE ELECTRIC utilities and independent grid managers that run interconnection simulations say they are simply following rules set by the Federal Energy Regulatory Commission (FERC), and their sole goal is to ensure a reliable supply of electricity.

FERC is now working on a revision of its rules. Renewable energy advocates, meanwhile, argue that some of the assumptions built into the simulations favor the business interests of old-style power companies, for whom the existing grid is working

PHOTO: MICKEY STRIDER/LOOP IMAGES/UNIVERSAL IMAGES GROUP/GETTY IMAGES

just fine. Those companies favor rules that protect their existing power plants from new competition, or shift the cost of valuable new transmission infrastructure to the new generators.

"There's definitely a political aspect, and a self-interest aspect, to all of this," Vander Vorst asserts.

Companies trying to build renewable power have won some victories. In 2021, they convinced SPP to drop an assumption that all wind and solar farms were simultaneously generating maximum power. Apart from being highly unlikely, this assumption produced simulations with so much transmission congestion that no new wind or solar proposals could even be considered in some areas. Now, SPP's simulations assume that previously approved wind and solar plants generate anywhere from zero to 75% of their capacity, depending on the exact scenario. The change came too late for Grady Martin, but the forecasts now have fewer overloaded power lines, and wind and solar projects are saddled with fewer upgrade costs.

Critics of the models, however, want revisions to go even further. Vander Vorst, for example, says interconnection studies should consider other ways to avoid overloaded power lines, apart from building new ones.

In real life, grid managers do this routinely, easing congestion on power lines by ordering some power plants to throttle down, while increasing power generation elsewhere on the grid to compensate. If new wind or solar farms create such problems frequently, Vander Vorst says, it's a sign that new transmission lines really are needed. But if it only happens rarely, the grid's gatekeepers could allow new projects to interconnect and just manage those problems if they occur.

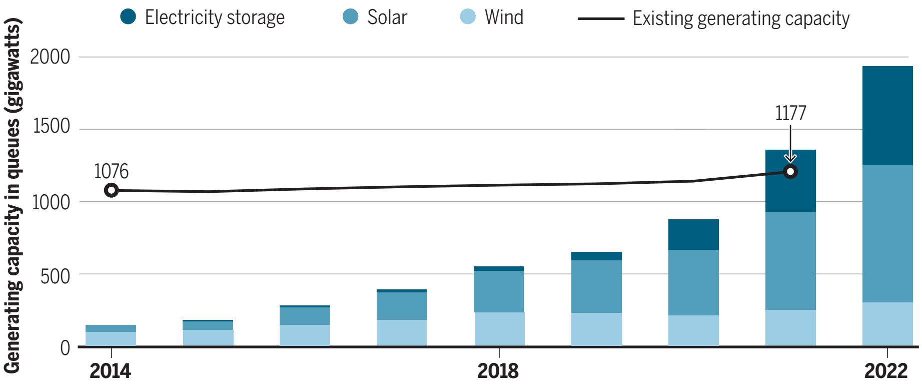
Texas already uses a version of this system. Solar and wind projects there face relatively few obstacles to interconnection. They do, however, face more financial risks once connected. If many projects connect to the grid in the same area, they can end up competing with each other for limited space on the transmission system, limiting the amount of power they can sell.

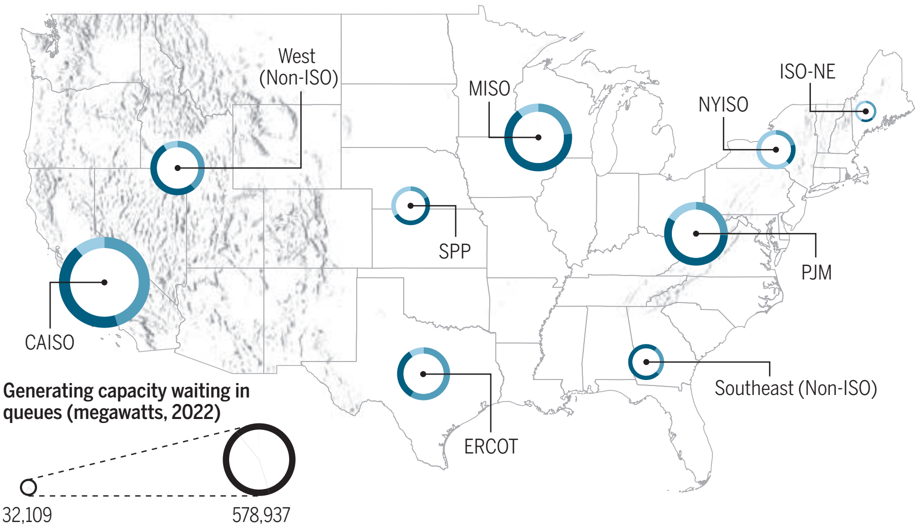
SPP isn't enamored of the Texas-style approach, Kelley says. His job, he says, is to make sure power plants can deliver their power to the grid with as few constraints as possible. "The last thing we want to do is to give the real-time operators a grid that they can't manage," he says.

Some renewable energy companies argue for another change: letting simulations incorporate new "grid-enhancing" technologies that allow the existing grid

Queued up

A rapidly increasing number of solar, wind, and energy storage projects are waiting to connect to the U.S. electricity grid (top). The generating capacity of these proposed projects exceeds that of the country's existing power plants. Many of the projects would be built in rural areas of the West and Midwest (bottom).





ISO is independent system operators, CAISO is California ISO, ERCOT is Electric Reliability Council of Texas, SPP is Southwest Power Pool, MISO is Midcontinent ISO, NYISO is New York ISO, ISO-NE is ISO New England, and PJM is Pennsylvania-New Jersey-Maryland. No federally reported data on existing generating capacity were available for 2022.



The U.S. National Renewable Energy Laboratory in Colorado has been studying ways to safely connect renewable electricity sources to regional transmission grids.

to carry more power than is currently assumed possible.

One technology is called dynamic line rating. It automatically adjusts the amount of current that power lines are allowed to carry based on local weather. On a cool, windy day, a transmission line often can handle substantially more power without overheating than it can on a hot, calm day. Related technologies are able to redirect the flow of electricity away from congested lines toward others that have excess capacity.

FERC has proposed rules that encourage grid operators to consider grid-enhancing technologies when studying interconnection requests. But many in the power industry are skeptical.

"Some of these things are really unproven, uncommercial kinds of technologies," Kelley says. "We can't just put something theoretical in the model and say: 'Oh, that makes the issue go away.'"

Dynamic line ratings, for instance, can be a kind of Band-Aid that helps control room operators resolve temporary congestion problems, Kelley says. But it's no replacement for new power lines. "You just typically need bigger pipes in order to move more energy," he says.

EVEN AS EXPERTS debate model tweaks that might ease clean energy connections, others are calling for more attention to the risk that wind and solar generators may react to electrical disturbances in unexpected ways. Those anomalies, they say, could pose a greater threat to the grid than congestion.

Researchers point to a near-disaster that occurred in Texas on 4 June 2022. It started with an equipment failure at a gas-burning power plant in Odessa, which caused the 550-megawatt plant to "trip," or go offline. The shutdown perturbed the voltage of electricity flowing across western Texas.

That should not have caused serious problems. Almost instantly, however, a dozen big solar farms in west Texas, which were generating 1700 megawatts at that moment, also tripped offline. The interruption briefly threatened to take down parts of the grid.

An investigation by the North American Electric Reliability Corporation (NERC), an industry group, pointed a finger at the solar plants' inverters—core equipment that converts the direct current from these generators into alternating current that matches what's already flowing on the grid. The inverters' programmable controls had been set in such a way that they didn't "ride through" the disturbance, as they are supposed to do.

The fix, which is still being implemented, involves reprogramming some of those software controls. Yet NERC and others say the Texas incident illustrates a broader vulnerability, rooted in the fact that wind and solar generators behave differently from traditional coal or gas plants.

Those old-style plants, with their massive rotating turbines, have an inherent physical ability to ride through disturbances. Wind turbines and solar panels, by contrast, rely entirely on automated, software-driven electronic systems to control their power output. "They do what they're told to do, and sometimes it's not the right thing," says Ian Hiskens, a power systems researcher at the University of Michigan.

In particular, these control systems are programmed to be "grid-following," meaning they track the 60-hertz cycles of voltage on the grid and deliver power that matches it precisely. But that can also leave them hypersensitive to intermittent disruptions in that signal, Hiskens says. Such disruptions can be caused by events as subtle as surges in power from a nearby wind farm as the wind picks up. "You can end up with interaction between two adjacent wind or solar farms, an action-reaction kind of thing," Hiskens says.

"You just typically need bigger pipes in order to move more energy."

David Kelley, Southwest Power Pool

These problems are more likely to pop up in parts of the grid that are far from old-style generators, which help maintain steady voltage and frequency—like conductors of an orchestra keeping everyone on tempo. At a distance, that tempo signal may be indistinct and hard to follow—a situation that's sometimes called a "weak grid."

In Australia, for example, Hiskens says authorities have noticed intermittent voltage oscillations in a part of the grid with lots of wind and solar farms. It hasn't been a major problem, he says, "but if you don't understand the oscillations, then you should be pretty wary. Because oscillations tend to be precursors to instability."

In the future, as wind and solar generators come to dominate the grid, some will likely be outfitted with "grid-forming" inverters that enable them to take over the tempo-setting role from older turbines. But that transition is still years away.

In the meantime, NERC would like to see the grid's gatekeepers add a new kind of model to their interconnection studies. These "electromechanical transient," or EMT, models can simulate the behavior of inverter control systems in exquisite detail, showing how they react to disturbances.

Grid operators such as SPP and the Midcontinent Independent System Operator are now carrying out such studies, but they are running into complications. The models require a lot of computing power, and people who know how to run them are in short supply. Perhaps more important, the models only deliver accurate results if they are using accurate information about the equipment that will be installed, and its control settings. That information is often unavailable, because when interconnection studies begin, solar or wind developers don't know what equipment they'll install years down the road.

"If you're going to do an EMT study in phase one [of interconnection], then you're high on crack," says Rhonda Peters, a consultant who's worked for many wind and solar developers. She says such studies make more sense if conducted closer to actual construction.

FERC IS CURRENTLY working on new regulations that could help ease the interconnection logjam. One set of proposed rules, released on 27 July, would impose fines on grid managers who take too long to consider interconnection requests. It also attempts to discourage renewable energy companies from clogging the queue with speculative projects that aren't likely to succeed.

In the short term, however, the interconnection problem is likely to worsen. The Inflation Reduction Act, a 2022 law that includes hefty financial incentives for wind and solar projects, is now fueling a new surge in applications. "There's so much activity in this industry, people are going bezonkers," Peters says.

Over the long term, studies suggest the most cost-efficient way out of the interconnection maze is for electric utilities to figure out in advance where new power lines are most needed—and then build them. Utilities, however, are typically reluctant to impose the rate increases needed to pay for new infrastructure. "The main barrier is still, who pays? And because they cannot solve it, that's where it all stops," says Julia Matevosyan, chief engineer for the nonprofit Energy Systems Integration Group. A rule FERC is now developing could help by shifting the burden of paying for new power lines, so new renewable energy projects won't have to shoulder the bulk of the cost.

For the moment, the interconnection simulations, for all their flaws, are actually revealing a fundamental truth. The U.S.'s energy transition will demand a far more ambitious expansion of electricity infrastructure than anyone, so far, has been willing to embrace. ■

Dan Charles is a journalist in Washington, D.C.

Chec
upda



VOLCANOLOGY

Anatomy of a volcanic eruption undersea

Submarine flows from the Hunga Tonga–Hunga Ha'apai eruption decimated seafloor cables

By **Rebecca Williams**[1] *and* **Pete Rowley**[2]

In December 2021, an undersea volcano in the southern Pacific Ocean, the Hunga Tonga–Hunga Ha'apai (hereafter called the Hunga volcano) began erupting. In January 2022 the eruption reached a powerful climax, triggering atmospheric waves that traveled around the globe and a tsunami that swept across the Pacific Ocean (*1, 2*). An estimated 75% of Earth's volcanoes are underwater, and 20% of all fatalities caused by volcanic eruptions since 1600 CE have been associated with underwater volcanism (*3*). Yet, explosive underwater eruptions are poorly understood. On page 1085 of this issue, Clare *et al.* (*4*) report that volcanic debris from the Hunga eruption traveled under the sea at an unprecedented distance and at record-breaking speed—more than 100 km, at velocities reaching 122 km/hour—and destroyed a vast network of seafloor telecommunication cables. Given that 95% of global communications are carried by seafloor cables, the findings highlight system vulnerabilities to underwater volcanism (*5*).

Clare *et al.* show that the particulate debris ejected from the Hunga volcano (the eruption column) collapsed vertically and directly into the ocean. This debris, consisting of rock and ash, then traveled as



A collapsing column of ejected ash and gas from the Hunga volcano caused damaging submarine currents in 2022.

volcaniclastic submarine density currents on the submerged slopes of the volcano. These were the fastest submarine density currents to have been recorded. The currents then traveled along the seafloor, destroying almost 200 km of cable that lay more than 15 km away from the volcano. Additionally, one of the cables was moved over 5 km by the currents. Bathymetric surveys, which map the shape and depth of underwater terrain, revealed 2-km-wide scours of the seabed where the currents eroded close to 100-m thickness of sea floor. The authors used the timing and extent of cable breaks, repeated bathymetric surveys, eruption observations, and rock core sampling to document the column collapse and resulting volcaniclastic submarine density currents. They calculated the velocity of the currents, determined the likely flow paths, and mapped where the currents eroded the seabed and deposited large volumes of volcanic debris on the seafloor.

Pyroclastic density currents typically form from eruption column collapse, which drives rapidly moving mixtures of volcanic ash and gas down the volcano's slopes. Clare *et al.* made the unexpected observation that Hunga's volcaniclastic submarine density currents transitioned between two types of flow behaviors, or rheologies. The morphology of the deposit from the currents and the occurrence of deposition on high-angle slopes of the volcano suggest that the currents must have initially been dense, granular, particle-laden, and carried by gas, similar to pyroclastic density currents on land. Characteristics of the deposits suggest that the currents then transitioned to water-carried, particle-laden submarine density currents, so-called turbidity currents. Understanding the internal dynamics of pyroclastic density currents is hampered by their destructiveness and opacity, which means not much can be revealed about them by long-distance observation. Much more is known about turbidity currents through direct monitoring, owing to new sensors and methods that provide measurements of, for example, flow velocity profiles (*6*).

Conceptual models of pyroclastic density currents have assumed that they are analogous to turbidity currents with respect to fluid mechanics. Moreover, theoretical work has suggested that pyroclastic density currents can propagate for substantial distances underwater without mixing with water (*7*). However, until now, evidence for this in the deep sea has been lacking (*8*). The switch from pyroclastic to turbidity current rheology, as observed by Clare *et al.*, indicates nuanced differences between having gas or water as the carrier fluid. This transition makes the Hunga eruption an ideal case study to further explore how analogous turbidity currents are to pyroclastic density currents. If the differences prove substantial, then some of the underlying assumptions of various frameworks used to model and understand pyroclastic density currents are up for challenge. This has implications for how volcanologists interpret the volcanic rock record, conduct hazard assessments at volcanoes, and use numerical models to calculate potential flow paths for future eruptions.

Recent studies have documented giant scours on the seafloor surrounding submarine volcanoes across the world, for example, at Macauley and Raoul Islands in the southwest Pacific (*9, 10*). Little is known about the eruptions that formed these scours, which likely occurred thousands of years ago. Conceptual modeling proposed that the scours formed during highly explosive eruptions, but this has not been tested. The seafloor erosion and scours reported by Clare *et al.* are comparable in size to those observed around other submarine volcanoes, which suggests that these morphological features of the seafloor are the result of powerful volcanic flows from submarine or near-shore volcanic eruptions. Thus, events with the magnitude of the Hunga eruption may not be uncommon. This highlights that large submarine volcanic eruptions are an underappreciated global risk. Effort must be invested in quantifying the dangers associated with these eruptions and exploring the engineering requirements for remediating them.

The findings presented by Clare *et. al.* contribute an important dataset that should inform new models that can guide the engineering and repair of submarine infrastructure. The Hunga eruption also raises questions about the largely unexplored hazard of submerged calderas and the direct collapse of the eruption column into water. This will spark research for many years to come, using the Hunga volcano as a natural laboratory to generate and test new hypotheses on explosive volcanism. Some of the areas to explore in the future include the seawater's role in driving or suppressing explosive eruptions and the effect of volcaniclastic submarine density currents on marine ecology. Some planetary scientists are even drawing analogies between the Hunga eruption and volcanoes on Mars (*11*). Ultimately, the Hunga volcano will be a vital case study for better understanding the risk that undersea and shallow-water volcanoes pose to the submarine environment and critical seafloor infrastructure. ■

[1]School of Environmental Sciences, University of Hull, Hull, UK. [2]School of Earth Sciences, University of Bristol, Bristol, UK. Email: rebecca.williams@hull.ac.uk

REFERENCES AND NOTES

1. C. J. Wright *et al.*, *Nature* **609**, 741 (2022).
2. S. J. Purkis *et al.*, *Sci. Adv.* **9**, eadf5493 (2023).
3. L. G. Mastin, J. B. Witter, *J. Volcanol. Geotherm. Res.* **97**, 195 (2000).
4. M. A. Clare *et al.*, *Science* **381**, 1085 (2023).
5. M. A. Clare *et al.*, *Earth Sci. Rev.* **237**, 104296 (2023).
6. P. J. Talling *et al.*, *Nat. Rev. Earth Environ.* (2023).
7. R. S. J. Sparks, H. Sigurdsson, S. N. Carey, *J. Volcanol. Geotherm. Res.* **7**, 97 (1980).
8. A. Freundt, J. C. Schindlbeck-Belo, S. Kutterolf, J. L. Hopkins, *Spec. Publ. Geol. Soc. Lond.* **520**, 595 (2023).
9. E. L. Pope *et al.*, *Earth Planet. Sci. Lett.* **493**, 12 (2018).
10. D. Casalbore *et al.*, *Sedimentology* **68**, 1400 (2021).
11. E. B. Hite *et al.*, *J. Volcanol. Geotherm. Res.* **401**, 106902 (2020).

10.1126/science.adk0181

MEDICINE

Tracking kidney transplant fitness

An implantable bioelectronic device detects the early signs of kidney transplant rejection in rats

By **Mohamad Zaidan**[1,2] *and* **Fadi G. Lakkis**[2]

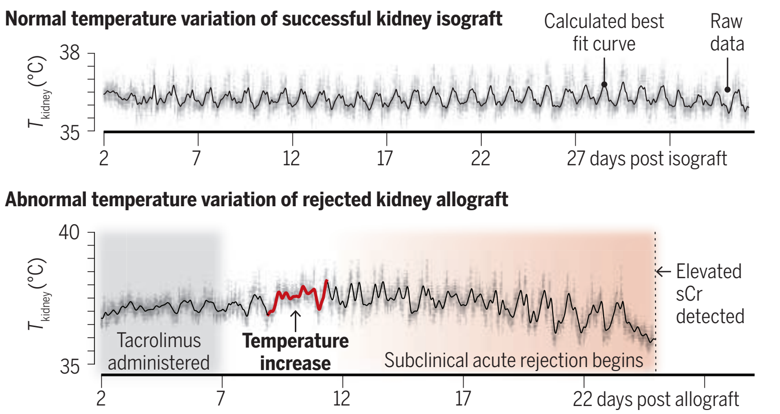
Kidney transplantation (allograft) is a life-saving treatment for patients with end-stage kidney disease. Despite considerable progress over the past decades, improving long-term kidney allograft survival remains a major challenge with 10-year survival rates in the US of ~55% for deceased donor transplants (*1*). Rejection, arising from the recipient's immune response to the allograft, accounted for ~30% of graft loss (*2*, *3*). Therefore, early detection and management of rejection are crucial to prevent irreversible kidney tissue damage and enhance patient outcomes. On page 1105 of this issue, Madhvapathy *et al.* (*4*) describe a fully implantable, wireless, bioelectronic system to detect subclinical acute rejection. Using rat kidney transplantation models, the authors established that monitoring kidney temperature (T_{kidney}) provides early warning of rejection with greater sensitivity than traditional biomarkers. This approach could uncover medication noncompliance and guide treatment to improve long-term graft outcomes.

Kidney biopsy is the "gold standard" for the diagnosis of transplant rejection. Nevertheless, patients and clinicians are reluctant to repeat such an invasive procedure throughout the transplanted organ's lifetime, hindering the possibility of regular monitoring of graft status. Routine biomarkers from blood and urine samples, such as serum creatinine, blood urea nitrogen, and proteinuria, lack the sensitivity and specificity required for the diagnosis of rejection. Notably, subclinical rejection, which corresponds to kidney allograft rejection with inflammatory lesions but stable renal function, occurs in up to 25% of patients, highlighting the lack of reliability of these markers to guide treatment decisions (*5*, *6*).

Alternative blood or urine biomarkers, including cell-free DNA (*7*), chemokine levels (*8*), and "omics" technologies, such as transcriptomics, proteomics, and metabolomics (*9*, *10*), have been developed to address these issues (*11*). However, these biomarkers still have drawbacks with specificity (positive predictive value), variability among transplant facilities, and the ability to discriminate between rejection types. Moreover, they only provide a snapshot of graft status, and using them to repeatedly monitor the graft over time may not be cost-effective (*12*). Crucially, none were designed to identify nonadherence to immunosuppressive drugs, which occurs in up to one-third of patients, is associated with increased risk of rejection, and is responsible for 15 to 20% of graft loss (*3*, *13*, *14*).

Kidney transplant monitoring

A wireless, bioelectronic device is implanted on transplanted kidneys in rats to continuously monitor kidney temperature (T_{kidney}) in real time. Increased T_{kidney} could be used as an early biomarker of local kidney inflammation, which is potentially related to subclinical acute rejection, before the rise in serum creatinine (sCr) levels.



Madhvapathy *et al.* constructed a bioelectronic device that is capable of continuous, real-time, and long-term monitoring of T_{kidney} and thermal conductivity (k_{kidney}), as surrogate markers for kidney inflammation and perfusion, respectively. The device consisted of a mechanically compliant biosensor (probe) connected through wires to an electronic module (receiver). Immediately after transplanting the rat kidney, the receiver was secured to the recipient's abdominal wall adjacent to the graft and the probe was mounted on the surface of the kidney under the kidney capsule. The device had no obvious effect on activity, food and water intake, or sleep-wake cycles. In the native kidney and in isografts (transplants from genetically identical rats), the physiological pattern of T_{kidney} was characterized by a periodic (1-day) circadian rhythm after an initial irregular period corresponding to postoperative recovery (0 to 2 days). Notably, whereas increased activity of the animal induced transient (<1 hour) variations in T_{kidney}, changes in ambient temperature did not. k_{kidney} increased initially after single kidney transplantation, reaching twice the value of an animal with both of its native kidneys intact. This demonstrates that the biosensor successfully detected physiological doubling of renal perfusion in an animal with a single kidney.

Madhvapathy *et al.* identified a distinctive pattern of T_{kidney} that provided a consistent and early warning sign of subclinical acute rejection. Unlike isografts, allografts in nonimmunosuppressed recipients showed an initial rise of T_{kidney}, but not k_{kidney}, followed by a decrease. This pattern was consistent across all rejecting kidney allografts and preceded any change in renal function by 3 days. In rats transiently treated with tacrolimus, a commonly used immunosuppressant, the change in T_{kidney} pattern that is indicative of rejection occurred 2 to 3 weeks before any change in serum biomarkers (see the figure). Moreover, tacrolimus administration was characterized by stabilization of the T_{kidney} fluctuations that no longer exhibited circadian rhythm, suggesting that the biosensor may have the added benefit of detecting whether immunosuppressive drugs are being taken.

The application of such an implantable device to humans would first require experiments in larger animals to ensure safety and adequate function. If successfully translated to humans, the device would represent a

[1]Assistance Publique des Hôpitaux de Paris, INSERM, Université Paris-Saclay, Hôpital de Bicêtre, Le Kremlin-Bicêtre, France. [2]Starzl Transplantation Institute, University of Pittsburgh, Pittsburgh, PA, USA. Email: mohamad.zaidan@aphp.fr; lakkisf@upmc.edu

considerable improvement over traditional biomarkers for patient and graft monitoring by allowing timely therapeutic interventions and the recognition of medication noncompliance. Beyond rejection, the recording of biophysical parameters, such as k_{kidney}, may also represent an opportunity to assess situations that affect graft blood flow, including hypovolemia (reduced blood volume) and hypotension (reduced systemic blood pressure), which may also lead to acute kidney injury.

The findings of Madhvapathy *et al.* are promising, but further evaluation is needed. The specificity of the changes observed in T_{kidney} and k_{kidney} for rejection should also be compared with alternative inflammatory conditions affecting the graft, such as pyelonephritis and BK virus–associated nephropathy. Additionally, it will be important to investigate whether continuous monitoring is also suitable for identifying borderline changes in the graft, which correspond to minimal inflammatory changes that may precede the occurrence of full-blown acute rejection or may be detrimental to long-term graft survival (*12*, *15*). Moreover, the ability of such a device to discriminate acute from chronic components of graft injury should also be explored because the models used by the authors are more consistent with acute kidney rejection leading to rapid graft loss. The potential combination of continuous monitoring with additional markers of kidney allograft injury would provide a more comprehensive assessment of the graft status, facilitating clinical decisions. It will also be important to consider the foreign-body response that may accompany the implantation of such a device. Similar to the response to pacemakers, vascular fibrosis and tissue encapsulation may limit the long-term application in humans. Although several hurdles remain to be overcome, the prospect of integrating continuous monitoring into clinical practice could represent a major step toward personalized organ transplant care. ■

REFERENCES AND NOTES

1. S. Hariharan, A. K. Israni, G. Danovitch, *N. Engl. J. Med.* **385**, 729 (2021).
2. Z. M. El-Zoghby *et al.*, *Am. J. Transplant.* **9**, 527 (2009).
3. M. Mayrdorfer *et al.*, *J. Am. Soc. Nephrol.* **32**, 1513 (2021).
4. S. R. Madhvapathy *et al.*, *Science* **381**, 1105 (2023).
5. A. Loupy *et al.*, *J. Am. Soc. Nephrol.* **26**, 1721 (2015).
6. R. B. Mehta *et al.*, *Kidney Int.* **102**, 1371 (2022).
7. L. Bu *et al.*, *Kidney Int.* **101**, 793 (2022).
8. C. Tinel *et al.*, *Am. J. Transplant.* **20**, 3462 (2020).
9. J. J. Friedewald *et al.*, *Am. J. Transplant.* **19**, 98 (2019).
10. W. Zhang *et al.*, *J. Am. Soc. Nephrol.* **30**, 1481 (2019).
11. C. M. Puttarajappa, R. B. Mehta, M. S. Roberts, K. J. Smith, S. Hariharan, *Am. J. Transplant.* **21**, 186 (2021).
12. M. Naesens, J. Friedewald, V. Mas, B. Kaplan, M. M. Abecassis, *Transplantation* **104**, 700 (2020).
13. I. Gandolfini, A. Palmisano, E. Fiaccadori, P. Cravedi, U. Maggiore, *Clin. Kidney J.* **15**, 1253 (2022).
14. A. Cherukuri *et al.*, *Kidney Int.* **96**, 202 (2019).
15. A. Cherukuri *et al.*, *Kidney Int.* **103**, 749 (2023).

10.1126/science.adj9517

GENOMICS

Chromosomal contacts change with age

The three-dimensional organization of the genome is remodeled throughout life

By **Schahram Akbarian**[1] *and* **Hyejung Won**[2]

Cell-specific transcriptomic and epigenomic programming during human brain development and differentiation are associated with dynamic changes in chromosomal conformations, which provide a structural foundation for a myriad of nuclear functions. These functions include the compartmentalization of the genome into intranuclear environments that are either facilitative (open compartments) or repressive of transcription, as well as the activation or fine-tuning of gene expression through promoter-enhancer looping (*1*, *2*). However, a fine-grained analysis of the three-dimensional (3D) organization of neuronal genomes at the single-cell level and across the life span has been missing, particularly for the human brain. On page 1112 of this issue Tan *et al.* (*3*) report that the chromosomes of some neuronal populations show potentially lifelong progressive changes in conformation in mice and in humans.

The 3D plasticity of the genome is thought to continue after the differentiation of neural progenitors to postmitotic neurons. For example, mapping of chromosomal conformation in neuron-enriched pools of brain nuclei from postnatal and young-adult mouse brain has shown a considerable capacity of the 3D genome to adapt in response to environmental enrichment (*4*) and during learning and memory (*5*, *6*). In addition, estrus cycle–driven chromosomal conformation dynamics have been observed in female mice (*7*).

Tan *et al.* used a DNA-DNA proximity assay called Dip-C and single-nucleus whole-genome amplification to study the dynamic landscape of chromosomal contacts in human and mouse brain. Their primary focus was cerebellar granule cells, which drive excitatory (glutamatergic) signaling in the cerebellum and are the most numerous neuronal subtype in the brain. The authors analyzed 5202 single nuclei from human cerebral and cerebellar cortex from 24 donors ranging in age from 5 weeks to 86 years and collected a similar dataset from mice. In terms of chromosomal conformation mapping, the size of these datasets is impressive. Tan *et al.* also performed standard single-nucleus transcriptome and chromatin accessibility profiling in the same cell type. The results of these analyses showed that cell differentiation–related changes in transcription (such as up-regulation of genes that encode synaptic proteins and ion channels or transcription factors that are critical for neuronal signaling) are, at least from a genome-wide perspective, completed at a relatively early stage in the life of these neurons. In humans, the entire population of cerebellar granule cells showed a mature transcriptome from the second year of postnatal life onward, which remained essentially unchanged thereafter.

By contrast, Tan *et al.* found evidence for an ongoing, perhaps even lifelong, reorganization of the chromosomal contact map, with successively increasing proportions of cerebellar granule cells displaying a fully matured 3D genome as the brain continued to age. This process extended into the seventh decade of life in humans and 12 months in mice. Tan *et al.* identified three specific elements of the 3D genome that seemed to undergo this progressive remodeling. One element was compartmentalization into facilitative and repressive chromosomal environments. Specifically, the authors identified a lifelong, age-progressive increase in open compartment scores (a metric for the degree by which a genomic locus is embedded into a facilitative chromatin environment) that encompass 100 to 200 cerebellar granule cell marker genes (see the figure). Paradoxically, these genes showed stable RNA expression levels from early childhood onward.

3D genome maturation in cerebellar granule cells was also defined by a steady rise in the proportion of intrachromosomal contacts that bypassed 10 to 100 Mb of sequence (ultra-long-range contacts), which increased from 19% at the most immature stage to 33% at the most mature stage in humans. Tan *et al.* also observed an increase in interchromosomal contacts associated with maturation. Therefore, increased compartmentalization in humans and mice could result from many long-range intra- and interchromosomal

[1]Icahn School of Medicine at Mount Sinai, New York, NY, USA. [2]Department of Genetics, University of North Carolina, Chapel Hill, NC, USA. Email: schahram.akbarian@mssm.edu; hyejung_won@med.unc.edu

contacts. Preliminary evidence from humans has indicated that this compartmentalization change also occurs in glia (*8*).

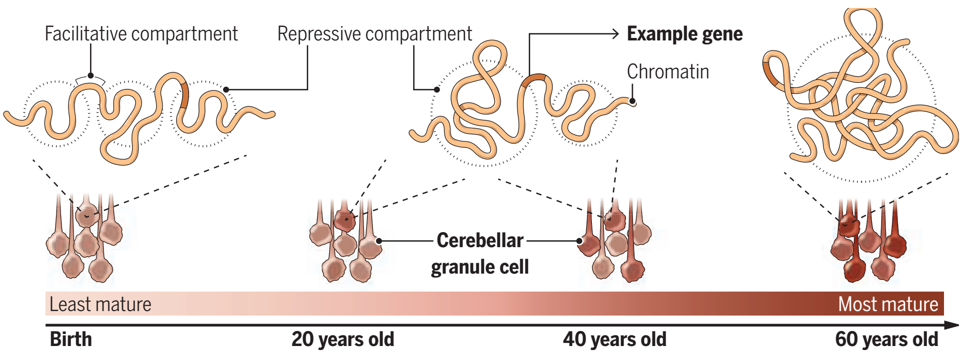
The functional implications of this finding are not yet clear. However, Tan *et al.* speculate that the changes that they saw in the 3D genome could be somehow related to the observation that cerebellar granule cells are extremely densely packed in large numbers in the cerebellar cortex, which requires each cerebellar granule cell to minimize energy expenditure and optimize use of space.

The extent of age-progressive 3D genome remodeling in cerebellar granule cells reported by Tan *et al.* is remarkable: A total of 539 1-Mb-wide sections of genome in human (473 in mouse) were involved, which is equivalent to twice the length of the largest human chromosome (chromosome 1). Furthermore, the increase in ultra-long-range chromosomal contacts during cerebellar granule cell maturation is particularly noteworthy because this effect was highly specific for this neuronal subpopulation. In the human and mouse brains studied by Tan *et al.,* the fraction of long-range conformations among all intrachromosomal contacts in other neuronal types, including cortical neurons and cerebellar Purkinje cells, was overall much lower (~10%) and increased by just 1% during maturation. The authors note that these differences among neuronal subtypes could be driven simply by differences in cell size, and thus nuclear size, with DNA-DNA proximity assays capturing longer-range contacts more frequently in cerebellar granule cell nuclei, which are among the smallest nuclei. However, this would still not explain the age-progressive increase in ultra-long-range intra- and interchromosomal contacts that was observed specifically in cerebellar granule cells.

Ultra-long-range interchromosomal contacts, such as those reported by Tan *et al.* that include those linking portions of mouse chromosomes 4 and 7, were previously reported for olfactory neuroepithelium (*9*) and mature cortical neurons. In the latter, these contacts were heavily regulated by repressive chromatin remodeling complexes (which assemble at specific genomic loci to silence transcription) that target active mouse retrotransposons (ancient retrovirus–derived DNA sequences) (*10*). It remains to be determined whether this mechanism contributes to the 3D genome dynamics observed in cerebellar granule cells. Regardless, Tan *et al.* have uncovered a genomic phenotype—the lifelong progressive remodeling of the chromosomal contact map—in a major neuronal cell type. Further research into the underlying mechanisms could deliver new insights about the genomics of aging, including, for example, a better understanding of retrotransposon silencing (*10*, *11*) and the recently described link between increased life span and reduced speed of RNA polymerase II–mediated transcription (*12*). ■

Changing chromosomal contacts

Although the transcriptome of human cerebellar granule cells is stable from around 2 years of age in humans, the three-dimensional organization of the genome in these neurons continues to mature well into adulthood. Mature cerebellar granule cells have more ultra-long-range intra- and interchromosomal contacts, with many gene loci shifting into chromatin compartments that are facilitative of transcription.



REFERENCES AND NOTES

1. H. Won *et al.*, *Nature* **538**, 523 (2016).
2. P. Rajarajan *et al.*, *Science* **362**, eaat4311 (2018).
3. L. Tan *et al.*, *Science* **381**, 1112 (2023).
4. S. Espeso-Gil *et al.*, *Front. Mol. Neurosci.* **14**, 664912 (2021).
5. J. Fernandez-Albert *et al.*, *Nat. Neurosci.* **22**, 1718 (2019).
6. A. Marco *et al.*, *Nat. Neurosci.* **23**, 1606 (2020).
7. D. Rocks *et al.*, *Nat. Commun.* **13**, 3438 (2022).
8. M. G. Heffel *et al.*, bioRxiv 10.1101/2022.10.07.511350 (2022).
9. L. Tan, D. Xing, N. Daley, X. S. Xie, *Nat. Struct. Mol. Biol.* **26**, 297 (2019).
10. S. Chandrasekaran *et al.*, *Nat. Commun.* **12**, 7243 (2021).
11. V. Gorbunova *et al.*, *Nature* **596**, 43 (2021).
12. C. Debès *et al.*, *Nature* **616**, 814 (2023).

10.1126/science.adk0961

IMMUNOLOGY

Macrophage superclusters to the rescue

Macrophage aggregates step in to capture blood-borne pathogens in the injured liver

By **Pieter A. Louwe**[1,2,3] *and* **Martin Guilliams**[1,2]

Positioned at the interface between intestinal and systemic circulation, the liver protects against blood-borne infections (*1*, *2*). To do this, the liver uses a dense network of sinusoids—narrow blood vessels that allow the slow flow of contaminated blood through the tissue. These vessels are populated by Kupffer cells—resident macrophages of the liver specializing in the capture of pathogens from circulation—that clean the blood before it reenters systemic circulation (*3*). However, in chronic liver disease and the associated fibrosis, this barrier system is perturbed, posing a serious risk of systemic dissemination of pathogens (*4*). On page 1066 of this issue, Peiseler *et al.* (*5*) report a striking feature of the fibrotic liver whereby macrophages outside the liver are recruited to the organ and form superclusters to clear blood-borne pathogens. In doing so, these superclusters retain the role of the liver as a filter, even when diseased.

In the healthy liver, Kupffer cells are embedded among stellate cells, hepatocytes, and sinusoidal endothelial cells, collectively referred to as the Kupffer cell niche. Together, these niche cells imprint liver-specific functionality on Kupffer cells (*6*), including the capacity to expand filipodia across sinusoids and engulf blood-borne pathogens through a process that involves a cell surface receptor, such as complement immunoglobulin receptor [(CRIg) also known as complement receptor V-set and immunoglobulin domain–containing protein 4 (VSIG4) (*3*)]. CRIg is expressed on human and mouse Kupffer cells, suggesting that CRIg-mediated capture of

[1]Laboratory of Myeloid Cell Biology in Tissue Homeostasis and Regeneration, VIB Center for Inflammation Research, Ghent, Belgium. [2]Department of Biomedical Molecular Biology, Faculty of Sciences, Ghent University, Ghent, Belgium. [3]Laboratory of Myeloid Cell Biology in Tissue Damage and Inflammation, VIB Center for Inflammation Research, Ghent, Belgium. Email: pieter.louwe@ugent.be; martin.guilliams@ugent.be

bacteria is an evolutionarily conserved function of these cells (*7*). In the steady state, the Kupffer cell compartment maintains itself, independent of circulating monocytes (*8*), but upon loss of Kupffer cells, such as during chronic liver injury, recruited monocytes repopulate the Kupffer cell pool (*6*, *9*). Because tissue specialization of monocytes into resident macrophages takes months and, once complete, seems resistant to change (*10*), imprinting of the tissue macrophage identity was thought to be irreversible.

Peiseler *et al.* report that in the fibrotic mouse liver, resident Kupffer cells dedifferentiated and lost many of their steady-state attributes, including the expression of CRIg. Thus, they effectively lost their liver-specific identity and specialized pathogen-capture capacity. Given the essential role of the niche cells, including stellate cells, in the imprinting of Kupffer cell identity, the authors propose that loss of contact area between Kupffer cells and stellate cells in the fibrotic liver underpins this loss of identity. Peiseler *et al.* show that Kupffer cells lost expression of the Kupffer cell identity master transcription factor liver X nuclear receptor alpha (*Nr1h3*) (*6*, *11*) in the fibrotic mouse liver. Expression of *Nr1h3* in Kupffer cells is induced by a combination of Notch signals from sinusoidal endothelial cells and bone morphogenic protein 9 (BMP9) [also known as growth differentiation factor 2 (GDF2)] signaling from stellate cells (*6*). This suggests that Notch-BMP9 signaling in Kupffer cells is altered in the fibrotic liver.

It remains unclear whether the down-regulated expression of *Nr1h3* in Kupffer cells and their subsequent loss of liver-specific identity are effects of fibrosis that directly limit contact between stellate cells and Kupffer cells, thus restricting their access to BMP9, or whether stellate cells and/or sinusoidal endothelial cells change their gene expression profile and stop producing the factors that imprint Kupffer cell identity. Notably, Peiseler *et al.* found that monocyte-derived Kupffer cells developing in the fibrotic liver were deficient in CRIg, which suggests that the Kupffer cell niche is broken in fibrosis.

Irrespective of the exact molecular mechanism at play, Peiseler *et al.* report that the mouse fibrotic liver was populated by Kupffer cells—both resident Kupffer cells and monocyte-derived Kupffer cells—that lack the expression of CRIg and the associated specialized pathogen-capture capacity. This highlights a limitation of the liver architecture to maintain pathogen-capture functionality when injured. However, the overall ability of the liver to trap blood-borne pathogens was only marginally attenuated.

To understand this discrepancy, Peiseler *et al.* investigated the location of macrophages in the mouse liver. In the healthy liver, Kupffer cells were exclusively situated close to the sinusoids, in line with previous observations (*6*). However, in the fibrotic liver, macrophages also formed large multicell aggregates within bigger venules and collateral blood vessels. These aggregates, or syncytia, comprise almost exclusively monocyte-derived macrophages. Syncytial macrophages had seemingly adopted liver-specific pathogen-capture machinery, as evidenced by high expression of CRIg, and captured blood-borne pathogens from circulation. In addition, stellate cells, which were sinusoid adjacent in the healthy liver, moved closer to venules and collateral blood vessels in the fibrotic liver (see the figure).

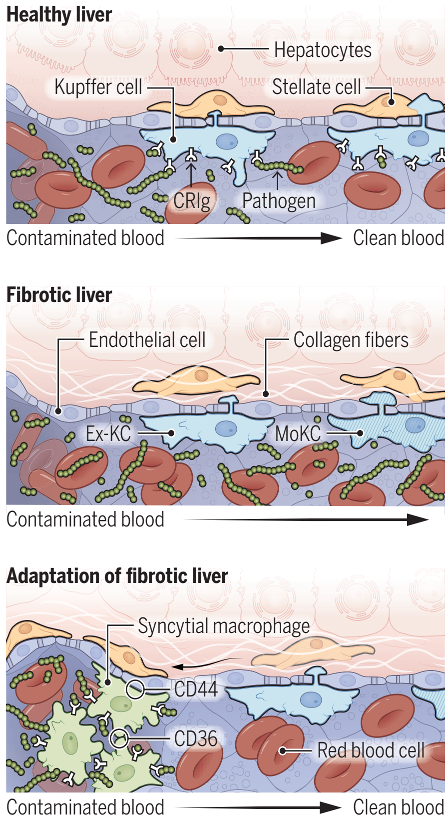
A hypothesis put forward by Peiseler *et al.* is that stellate cells migrate to imprint pathogen-capture functionality onto syncytial macrophages. However, a reverse mechanism, whereby the migration of stellate cells precedes syncytia formation, could also explain the findings. If so, stellate cell migration during early disease could drive the loss of functionality in the sinusoid-adjacent Kupffer cells. Regardless, these findings raise the question of whether the liver niche responds in a coordinated fashion to injury or if, upon injury, various niche cells respond and migrate in a disjointed manner, thus breaking the steady-state niche and potentially establishing a new inflammatory niche.

Peiseler *et al.* also identified two adhesion molecules present on syncytial macrophages—CD44 and platelet glycoprotein 4 (CD36)—as key players in the formation of superclusters. Hyalyronan, the ligand for CD44, was highly expressed on the fibrotic liver venule and collateral wall, thus allowing adhesion of CD44-expressing monocytes that differentiate into presyncytia macrophages. CD36 was dispensable for this early adhesion but was required for monocytes and subsequently macrophages to fuse into syncytia. Notably, recruited macrophages deficient in CD36 failed to form syncytia or capture blood-borne pathogens despite increased expression of CRIg. Hence, the effective uptake of pathogens by syncytia seems to be an interplay between correctly primed macrophages and the formation of macrophage superclusters. The long-term fate of the syncytial macrophages is unknown. Indeed, whether full anatomical and functional liver recovery is possible or if these syncytia present a lasting change is unclear.

The observations of Peiseler *et al.* reveal a notable adaptation of the injured mouse liver to maintain its capacity to clear circulating pathogens. The authors also identified macrophage syncytia in the diseased human liver, which suggests an evolutionarily conserved mechanism. Whether tissue-resident macrophage identity and functionality in any tissue are reversible in the context of tissue inflammation—as seems to be the case for the liver's Kupffer cells—is not yet clear. If so, such dedifferentiation is likely to affect tissue-resident macrophages in various injury settings and could have great clinical relevance. ■

Adaptation of the fibrotic liver

In the healthy liver (top), Kupffer cells are imprinted by stellate cells, endothelial cells, and hepatocytes to capture blood-borne pathogens using complement immunoglobulin receptor (CRIg). In the fibrotic liver, imprinting fails, resulting in ex-Kupffer cells (ex-KCs) and monocyte-derived Kupffer cells (moKCs) that lack CRIg and consequently fail to engulf pathogens (middle). The liver adapts (bottom) by forming clusters or synctia of CRIg-expressing macrophages, thereby retaining filtering capacity. These macrophages attach to each other with platelet glycoprotein 4 (CD36) and attach to venule or collateral endothelial cells with CD44.



GRAPHIC: N. BURGESS/*SCIENCE*

REFERENCES AND NOTES

1. C. N. Jenne, P. Kubes, *Nat. Immunol.* **14**, 996 (2013).
2. M. L. Balmer *et al.*, *Sci. Transl. Med.* **6**, 237ra66 (2014).
3. Z. Zeng *et al.*, *Cell Host Microbe* **20**, 99 (2016).
4. M. Bhattacharya, P. Ramachandran, *Nat. Immunol.* 10.1038/s41590-023-01551-9 (2023).
5. M. Peiseler *et al.*, *Science* **381**, eabq5202 (2023).
6. J. Bonnardel *et al.*, *Immunity* **51**, 638 (2019).
7. M. Guilliams *et al.*, *Cell* **185**, 379 (2022).
8. M. Guilliams, C. L. Scott, *Immunity* **55**, 1515 (2022).
9. C. Zwicker *et al.*, *Front. Immunol.* **12**, 690813 (2021).
10. Y. Lavin *et al.*, *Cell* **159**, 1312 (2014).
11. M. Sakai *et al.*, *Immunity* **51**, 655 (2019).

10.1126/science.adj9725

RETROSPECTIVE

Evelyn M. Witkin (1921–2023)

Pioneer of cell mutagenesis and DNA repair research

By **Joann B. Sweasy**[1] and **Philip C. Hanawalt**[2]

Evelyn M. Witkin, renowned geneticist, died on 8 July. She was 102. Together with Miroslav Radman, Evelyn discovered the first coordinated cellular stress response to DNA damage, called the SOS response. Her research pioneered the fields of DNA repair and mechanisms of cellular mutagenesis. Her foundational discoveries have had major impacts in cancer research and advanced the study of drug resistance and the development of the immune repertoire.

Born on 9 March 1921 in New York City, Evelyn Ruth Maisel received a bachelor's in biology from New York University in 1941 and married Herman A. Witkin in 1943. During her graduate studies at Columbia University, under the mentorship of Theodosius Dobzhansky and Salvador Luria, she spent summers conducting research at Cold Spring Harbor Laboratory on Long Island. After completing her PhD in genetics in 1947, she returned to Cold Spring Harbor Laboratory as a postdoctoral fellow and staff member. In 1955, Evelyn joined the faculty of the State University of New York Downstate Medical Center in Brooklyn. She then moved to Rutgers University, first Douglass College and subsequently the Waksman Institute of Microbiology, where she was professor of genetics from 1971 until her retirement in 1991.

Evelyn had planned to study genetics in graduate school, using fruit flies (*Drosophila melanogaster*) as a model organism. However, upon learning of Salvador Luria and Max Delbrück's use of the bacterium *Escherichia coli* in their groundbreaking discovery that mutation is random, she decided to focus her research on bacteria, with the goal of understanding the nature of the gene. In her first experiment as a graduate student, Evelyn set out to generate mutations by treating *E. coli* strain B with ultraviolet light (UV). There was no previous information on the UV sensitivity of *E. coli*, so she selected a relatively high dose, which killed all but four colonies of cells. Instead of discarding the results and starting again, Evelyn hypothesized, and later proved, that these colonies arose from mutants. She named one of the colonies "*E. coli* B/r" (*E. coli* B resistant to radiation) and demonstrated that the *E. coli* B/r strain was resistant to both UV and x-rays. Microscopic observation revealed that, unlike *E. coli* B, the B/r cells could still divide when treated with a high dose of UV. Evelyn concluded that the ability to stop cell division is directly related to the lethal effects of UV.

During her years at Cold Spring Harbor Laboratory, Evelyn developed a close relationship with geneticist Barbara McClintock. McClintock's discovery that



genes move from one location to another, through the rapid joining of chromosome ends, led Evelyn to think about DNA repair events within chromosomes. She pioneered the idea that, in addition to photoreactivation, there must be a DNA repair process that operates in the dark. One of us (P.C.H.), intrigued by Evelyn's insights, engaged her in productive discussions beginning in the early 1960s. Evelyn was particularly excited when the pathway of transcription-coupled DNA repair was revealed some years later.

In a landmark paper in 1967, Evelyn hypothesized that bacterial cell division is controlled by a repressor that, like the lambda bacteriophage repressor, is inactivated by a complex process initiated by the presence of replication-blocking lesions in DNA. She further suggested that this might not be the only cellular function to exhibit induction by DNA damage.

In 1971, Miroslav Radman, at that time a postdoctoral fellow at Harvard, proposed that UV induces a mutagenic mode of DNA replication. Thus was born the SOS hypothesis, named in reference to the Morse code distress call. Evelyn subsequently showed that the *lexA* and *recA* genes are required for the SOS response to DNA damage and that RecA protein plays an important direct role in SOS mutagenesis. We now know that many genes are induced as part of the SOS response, including the activity of an error-prone DNA polymerase, Pol V mut, which is activated by RecA protein. Evelyn also characterized the classic phenomenon called mutation frequency decline (MFD), which entails the rapid and irreversible decrease in UV-induced mutations that arise while protein synthesis is inhibited. Notably, the *mfd* gene product was later shown to be an essential factor for transcription-coupled excision repair.

Evelyn worked in the laboratory throughout her career, preferring research and mentoring to administration. One of the highlights of her day was opening the incubator door, removing the agar plates, and looking at the resulting bacterial colony counts from the latest experiment. One of us (J.B.S.) shared a bench with Evelyn and often benefited from her willingness to discuss her results and interpretations of them in detail. She taught her mentees how to think clearly about their results and emphasized that the simplest conclusion was usually the correct one. She also gave them room to mature as scientists. A believer in the importance of accurate and articulate writing, she presented a copy of Strunk and White's *The Elements of Style* to each student when they joined her lab.

In recognition of her seminal contributions to the fields of mutagenesis and DNA repair, Evelyn was elected to the US National Academy of Sciences in 1977. Her many awards included the 2000 Thomas Hunt Morgan Medal, the 2002 National Medal of Science, and the 2015 Albert Lasker Basic Medical Research Award. Beyond her scientific pursuits, she served as director-at-large of the New York Browning Society and studied how Charles Darwin and poet Robert Browning, who were contemporaries, may have influenced each other.

Evelyn was an icon in the field of genetics, a person of immense integrity and generosity, and a dear friend. Her palpable joy and enthusiasm for science along with her extraordinary intellectual discussions of findings created an encouraging environment that facilitated the development of her students into outstanding independent scientists, who now stand ready to carry on her legacy. ■

[1]The University of Arizona Comprehensive Cancer Center, Tucson, AZ, USA. [2]Department of Biology, Stanford University, Stanford, CA, USA. Email: jsweasy@arizona.edu; hanawalt@stanford.edu

10.1126/science.adk1028

POLICY FORUM

CORPORATE ACCOUNTABILITY

Hold big business to task on ecosystem restoration

Corporate reporting must embrace holistic, scientific principles

By **Timothy A. C. Lamont[1], Jos Barlow[1], Jan Bebbington[2], Thomas Cuckston[3], Rili Djohani[4], Rachael Garrett[5,6], Holly P. Jones[7,8], Tries B. Razak[9], Nicholas A. J. Graham[1]**

Large transnational corporations (TNCs) could use their considerable finances, labor, manufacturing infrastructure, and logistics expertise to play key roles in upscaling ecosystem restoration efforts, which are vital for achieving global biodiversity, climate, and development targets (*1*, *2*). Many TNCs are positioning themselves as environmental leaders, carrying out restoration that goes far beyond legal obligations to offset their own environmental impacts. This promise of corporate-led progress is alluring, and has delivered benefits in some cases, but is also fraught with risks. Well-intentioned efforts can do more harm than good (*3*), and some corporations oversell their efforts for reputational enhancement (greenwashing). Our evaluation of sustainability reports of 100 of the world's largest businesses reveals the extent to which TNCs are claiming to contribute to—but failing to report on—ecosystem restoration. Increased rigor, consistency, transparency, and accountability are needed to ensure that corporate-led restoration delivers quantifiable, beneficial, and equitable outcomes.

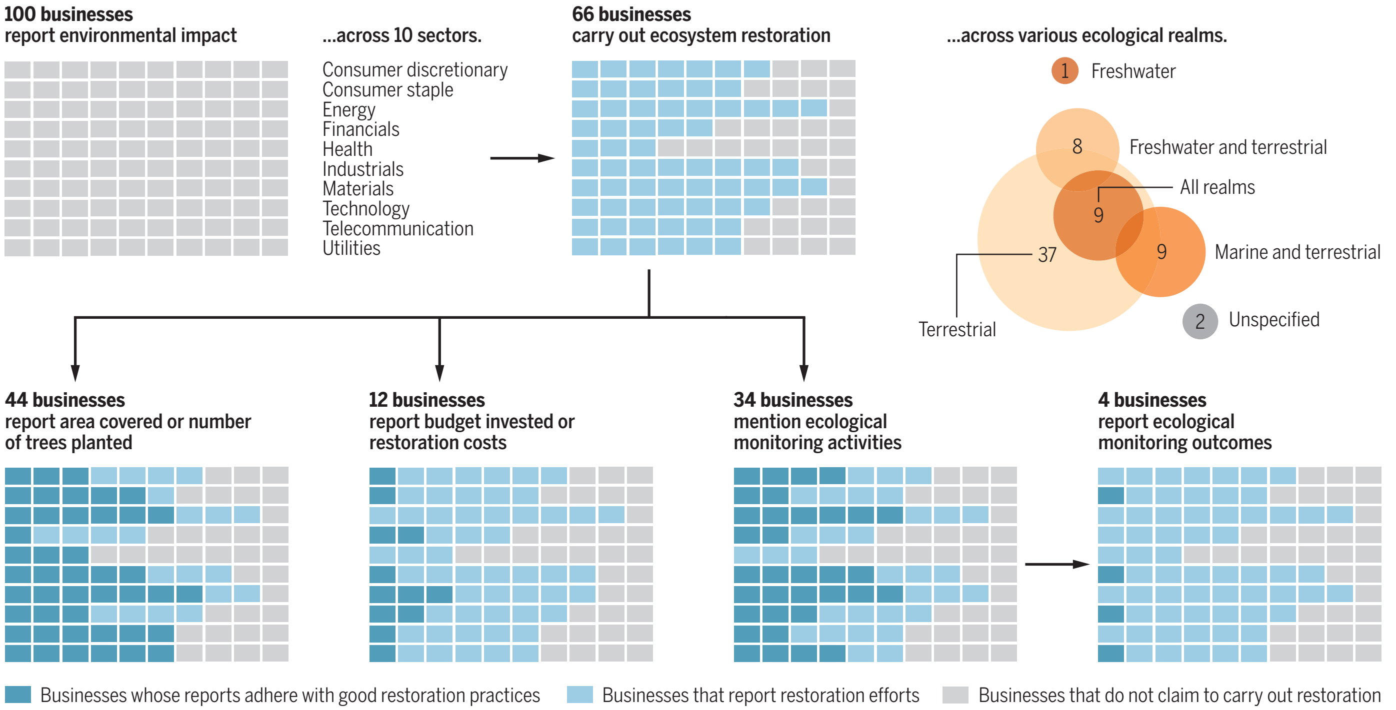
CURRENT STATE OF PLAY

In recent decades, large corporations have increasingly articulated and reported on their environmental responsibilities (*4*). Much of this progress stems from legal mandates for corporations to mitigate or offset their direct operational impacts. For example, the 2022 Kunming-Montreal Global Biodiversity Framework outlines that governments must implement requirements for large corporations to "regularly monitor, assess, and transparently disclose their risks, dependencies, and impacts on biodiversity." This reflects existing country-based requirements for businesses to conduct Environmental Impact Assessments (EIAs) to quantify and reduce their environmental damage. Ecosystem restoration that offsets damage is often reported under such frameworks.

Additionally, beyond these efforts and reporting required by law, businesses collectively have voluntarily pledged to, for example, plant billions of trees (the Trillion Trees Corporate Alliance), hundreds of thousands of corals (the Hope Grows program) and tens of thousands of mangroves (the Million Mangroves program). Private-sector reporting initiatives are emerging to encourage companies to voluntarily measure and disclose their biodiversity impacts more broadly, beyond legal requisites. These reporting initiatives include the Global Reporting Initiative (GRI), the Taskforce on Nature-related Financial Disclosures (TNFD), the Science Based

Reporting of ecosystem restoration by 100 of the world's largest businesses

Each block represents one business, chosen as the ten biggest corporations (by revenue) in each of the ten most represented industrial sectors



GRAPHIC: D. AN-PHAM/*SCIENCE*

Targets Network (SBTN), the International Sustainability Standards Board, and the Align project. These standards—although voluntary in theory—are often effective tools for regulating TNCs because they generate normative ideas about "good behavior" among large companies (*5*). Such guidelines have helped increase the rigor of reporting across a range of sustainability metrics in an era when many businesses are attempting to shift their public identities toward environmental stewardship. Amid this general pattern of heightened reporting, however, specific and important details on emerging attempts by TNCs to enter the wider ecosystem restoration agenda lag worryingly behind.

Our analysis of sustainability reports from 100 of the world's largest corporations (see supplementary materials) reveals that two-thirds of these companies state that they carry out various forms of restoration. These efforts encompass tree planting, repairing recent specific environmental damage, and assisting long-term recovery of extensively degraded ecosystems. Particularly active participation is observed in the energy and materials sectors, in which 9 of the top 10 companies describe restoration efforts. Concerningly, however, across all sectors there is a marked lack of rigor in defining restoration, outlining methods, and quantifying outcomes (see the figure). Most reports do not differentiate between projects that simply mitigate operational impacts of firms (as legally required) and those going beyond this to contribute toward wider global restoration goals. A third of reports fail to mention the size of any of their restoration projects, nearly 80% provide no information about financial costs, and more than 90% fail to report a single ecological outcome. None of the 100 reports quantify social or economic impacts on local stakeholders or traditional owners. This near-total lack of transparency in both ecological and socioeconomic reporting means that there is no way to quantify the amount of restoration being done or to confirm that its outcomes are indeed beneficial. Put simply, the evidence base supporting large corporations' claims about ecosystem restoration is wholly insufficient.

KEY PRINCIPLES

Substantial improvements in accountability and evaluation of progress could be achieved if corporate reporting was structured around key principles from restoration science. Existing sustainability guidelines may help to drive and improve this process, but they must be implemented in ways that recognize the need for holistic evaluation of multiple environmental and socioeconomic metrics. For example, the GRI Standards are used by more than two-thirds of the world's large corporations (*6*), defining the information that should be provided in business-relevant sustainability reporting. Restoration currently falls within the ambit of GRI Standard 304 for biodiversity (*7*), which is under revision after a period of public consultation. In the proposed standard, restoration-specific disclosure requirements are focused on mitigating business operational impacts and limited to documenting the spatial area of projects (GRI 304–5, a, iii). This guidance lacks the stringency and detail required to appropriately report on the increasingly ambitious restoration activities carried out by many corporations. Although some existing reporting requirements are relevant to certain aspects of restoration (for example, GRI 304-6, on reversing biodiversity loss; GRI 305-5, on offsetting carbon emissions; and GRI 413-1, on engaging local communities), the current enthusiasm and ambition surrounding large-scale corporate restoration demands more explicit, restoration-specific guidance.

Considerable work in the past 5 years has identified a set of complementary principles that lead to positive restoration outcomes (*3*, *8*–*10*). These principles each pertain to different stages of the restoration process (see the table). If companies' reporting gave details about each of these principles, it would allow better assessment of the scale and impacts of restoration. Improved reporting of these principles would also likely drive improvements across the different phases of restoration; companies with holistic reporting will be further incentivized to improve the design, implementation, and measurement of their projects. Notably, each of the seven principles was reported appropriately by at least three corporations (see the table), demonstrating that it is possible to provide compliant evidence within a standard reporting framework. However, these examples were rare; most principles were reported on by only a handful of corporations, and no single corporation reported appropriately on all seven principles. Further, most examples of compliant reporting were given for just one case study within a corporation's portfolio of restoration projects, rather than comprehensive reporting that covers all restoration activity.

IMPLEMENTING TRANSPARENCY

It is clear that existing corporate reporting fails to adequately communicate the outcomes of different restoration activities. International guidelines and national laws are currently designed to support biodiversity impact reporting and generally consider restoration as a tool to compensate for direct operational impacts. However, this does not account for efforts of corporations that go beyond this "bare minimum" and attempt to drive net-positive global change.

International guidelines (such as GRI, SBTN, and TNFD) must provide a new framework for restoration activities that are additional voluntary contributions. As the details of this new reporting framework emerge, trade-offs are likely to develop between the depth of reporting and the cost involved in providing information. As such, a key challenge will be to define "minimum standards" associated with each restoration principle, that ensure the provision of necessary data without excessive cost. For example, minimum standards on proportionality might require information about the spatial extent and number of organisms planted in each individual restoration project, and minimum standards on permanence might require a statement of the number of years committed to maintenance, monitoring, and reporting. Standards must also be flexible enough to accommodate diverse targets; for example, projects aiming to preserve specific endemic species will report different outcomes to those focused on restoring ecosystem services. As such, guidelines should encourage corporations to engage with local stakeholders, traditional owners, and subject-specific experts to apply these principle-based standards in the individual context of each specific project (*11*).

National or jurisdictional permitting authorities can align with these guidelines by requiring corporations to report specifically about whether individual restoration projects are legally required offsets or additional voluntary efforts. Summative reporting should be required in both cases, but with different approaches and goals. Legally required offsets should already be

[1]Lancaster Environment Centre, Lancaster University, Lancaster, UK. [2]Pentland Centre, Lancaster University Management School, Lancaster, UK. [3]Department of Accounting, Birmingham Business School, University of Birmingham, Birmingham, UK. [4]Coral Triangle Center, Sanur, Bali, Indonesia. [5]Environmental Policy Lab, ETH Zürich, Department of Humanities, Social and Political Sciences, Zürich, Switzerland. [6]Department of Geography and Conservation Research Institute, Cambridge University, Cambridge, UK. [7]Department of Biological Sciences, Northern Illinois University, DeKalb, IL, USA. [8]Institute for the Study of the Environment, Sustainability and Energy, Northern Illinois University, DeKalb, IL, USA. [9]Research Centre for Oceanography, National Research and Innovation Agency (BRIN), Jakarta 14430, Indonesia. Email: tim.lamont@lancaster.ac.uk

Compliance of corporate reporting with seven principles of restoration best practice

Of 100 corporations studied, 66 reported that they carry out restoration. Analysis of adherence to restoration principles is limited to those 66 corporations. Principles are grouped by the phase of restoration activity to which they are most relevant. See table S1 for further details.

RESTORATION PRINCIPLE	EXPLANATION AND POSSIBLE REPORTING METRICS	PERCENTAGE OF CORPORATIONS COMPLIANT (n = 66)
Most relevant to project design		
Mitigation hierarchy	Restoring degraded habitat is generally less effective than conserving intact ecosystems. Corporations should work within the "mitigation hierarchy" (*13*) by reporting their efforts to conserve existing habitat as a precursor to planning restoration.	39% (26)
Inclusive governance	Corporations should work with local stakeholders and decision-makers during planning and implementation of restoration projects. Reports should describe these partnerships with a focus on empowering local communities and traditional owners.	21% (14)
Most relevant to project implementation		
Permanence	Projects should plan for lasting impact. Reports should state the number of years committed to maintenance and monitoring, and/or survival rates and durations of previous projects.	11% (7)
Proportionality	Restoration should be proportional to environmental damage. Reports should disclose the area restored and/or budget invested in restoration, to demonstrate the extent of corporations' work.	70% (46)
Most relevant to measurement and reporting of outcomes		
Monitoring and transparency	Corporations claiming to restore ecosystems should prove that their initiatives are having desired ecological impact. Projects should define specific goals of restoration and regularly monitor their progress against these goals using quantitative ecological data. They should publish results in open access reports.	6% (4)
External benefits	Projects should target and report benefits beyond ecosystem recovery. For example, restoration can support local livelihoods, community engagement, education, research, training, and capacity building.	21% (14)
Reference ecosystems	In many cases, historic baselines are no longer feasible owing to changing environmental conditions. Projects should monitor local "reference ecosystems" to guide their efforts in restoring locally appropriate species compositions that are resilient to current and emerging threats.	5% (3)

covered by EIA frameworks and submitted to appropriate regulatory bodies for audit. However, additional voluntary projects should outline a separate set of publicly accessible aims, plans, and reports. Plans should be evaluated by experts as a prerequisite to activities commencing, to ensure that they avoid social harms, are ecologically viable, and have clear reporting structures in place (*12*).

Reporting on nonmandatory activities would be of interest to an array of corporate report users. This includes company owners (especially those with ethical screening in place), funders (usually banks), information providers to capital markets (such as corporate rating agencies), peer-group companies (especially those in the same industry or geographic region), and other stakeholders who wish to understand the outcomes of corporate restoration (such as nongovernmental organizations and scientists). These cases all provide incentives for regulatory bodies to ensure that this information is in the public domain.

Importantly, the reporting and evaluation of such projects must acknowledge that restoration is difficult and even well-planned projects sometimes fail or require changes in strategy. The aim of this exercise should therefore not be to penalize or publicly shame projects that fall short of original targets but to summarize current progress, promote learning from mistakes, and encourage improvements in practice.

RESTORATION AT A CROSSROADS

Consistent reporting is required to accurately assess the impact of big business' contributions to the United Nations Decade on Ecosystem Restoration, now in its third year. This increased transparency should also be in the interests of large corporations themselves, who stand to gain much from demonstrating to their shareholders, employees, consumers, and regulators that they make meaningful contributions to global sustainability. Businesses already have strong incentives to demonstrate the value of their restoration initiatives; strengthened reporting frameworks can help to deliver this accountability more effectively.

If managed carefully with transparent reporting, the world's largest corporations could considerably upscale ecosystem restoration, provide an evidence base from which others may learn, and garner public recognition for their efforts. Conversely, inadequate reporting by these organizations undermines accountability in ways that could threaten the credibility of the global restoration movement, exacerbate environmental damage, and create social injustice. Corporate involvement will certainly transform the future of ecosystem restoration; now, policy interventions must determine whether that change is for better or worse. ■

REFERENCES AND NOTES

1. B. B. N. Strassburg *et al.*, *Nature* **586**, 724 (2020).
2. P. Kareiva, M. Marvier, *Bioscience* **62**, 962 (2012).
3. S. Löfqvist *et al.*, *Bioscience* **73**, 134 (2022).
4. C. Antonini, C. Larrinaga, *Sustain. Dev.* **25**, 123 (2017).
5. J. Bebbington, E. A. Kirk, C. Larrinaga, *Account. Organ. Soc.* **37**, 78 (2012).
6. J. McCalla-Leacy, J. Shulman, R. Threlfall, "Big shifts, small steps: Survey of sustainability reporting 2022" (KPMG, 2022); https://home.kpmg/xx/en/home/insights/2022/09/survey-of-sustainability-reporting-2022.html.
7. GRI, "GRI 304: Biodiversity" (GRI, 2016); https://www.globalreporting.org/standards.
8. G. D. Gann *et al.*, *Restor. Ecol.* **27**, S1 (2019).
9. T. Osborne *et al.*, *Glob. Environ. Change* **70**, 102320 (2021).
10. A. Di Sacco *et al.*, *Glob. Change Biol.* **27**, 1328 (2021).
11. K. Schumacher, *APO Productivity Insights* 10.2139/ssrn.4303609 (2022).
12. K. Schumacher, *Nature* **584**, 524 (2020).
13. J. Ekstrom, L. Bennun, R. Mitchell, *A Cross-Sector Guide for Implementing the Mitigation Hierarchy* (Cross Sector Biodiversity Initiative, 2015); http://www.csbi.org.uk/our-work/mitigation-hierarchy-guide.

ACKNOWLEDGMENTS

This research was funded by an 1851 Royal Commission Research Fellowship (to T.A.C.L.), a Royal Society University Research Fellowship (URF\R\201029 to N.A.J.G.), and UKRI Natural Environment Research Council grant NE/X015262/1 (to J.Ba.). Data supporting this study are available as supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh2610

10.1126/science.adh2610



BOOKS *et al.*

The rejection of COVID-19 precautions is illustrat of broader anti-scientific attitudes, argues the author.

PUBLIC HEALTH

The legacy of Covid denial

A physician warns of the broader implications of pandemic backlash

By **Arthur Caplan**

A controversy emerged at the conclusion of the Vietnam War over who, if anyone, would write the "definitive" book on that conflict. Many works, both fiction and nonfiction, appeared over the years that followed, but none seemed worthy of the appellation. That will not be true for the COVID-19 pandemic. Virologist and vaccine expert Peter Hotez has provided the definitive work on this era with his new tome, *The Deadly Rise of Anti-science*, which presents a short, readable overview of the lives lost to the virus and the fierce backlash against sensible public health measures perpetuated by right-wing politicians and media outlets, by uninformed citizens and bad actors on social media, and by a small gaggle of contrarian doctors and scientists who irresponsibly engaged in various degrees of Covid denialism.

Hotez is at his best when he is describing the preventable COVID-19 deaths that followed the abandonment of vaccine and mask mandates and the accompanying decision, made by many, not to be vaccinated against COVID-19 in the second half of 2021 and early 2022. He estimates that ~200,000 unvaccinated people died unnecessarily in the United States, many in red states, including Texas, where Hotez lives. This number is consistent with a recent scholarly analysis conducted by the National Bureau of Economic Research (*1*).

The reviewer is at the Division of Medical Ethics, NYU Grossman School of Medicine, New York, NY 10016, USA. Email: arthur.caplan@nyumc.org

As anti-vax voices grew louder, and the Biden administration bailed on its own effort to promote boosters, Hotez joined a cadre of other physician-scientists, including Anthony Fauci and Paul Offit, to try to combat the spread of vaccine disinformation. Public health has traditionally relied on a moral framework that targets groups and communities, presuming that individuals will respond to advice and requirements out of a sense of duty to care for and protect others. COVID-19 arrived as a very different ethic took hold in much of the United States, one that privileged liberty and individualistic self-regard. Public health and its spokespeople thus became demonized, in part because they bore a moral message of community danger and the need for sacrifice that an increasingly divided, polarized, and egotistical society did not want to hear.



The Deadly Rise of Anti-science: A Scientist's Warning
Peter J. Hotez
Johns Hopkins University Press, 2023. 240 pp.

Hotez, a man whose compassion and dedication to healing are on display throughout the book, describes the personal cost he paid for his efforts to publicly combat anti-scientific policies and rhetoric during the pandemic when he devoted huge amounts of time to communicating his evidence-based opinion to an increasingly polarized public. The toll included doxing, death threats, intimidation from government officials and hate groups, and harassment of his family and some of his colleagues. A common refrain from his ideological opponents was that he was a "pharma shill," as if anti-vaxxers and Covid deniers were not making fortunes by selling phony treatments and preventatives.

In later chapters, Hotez examines the legacy of Covid denial as it morphed into an aggressive, well-organized, and well-funded attack on mainstream science. Such efforts have been exacerbated by neofascist hate groups proudly flaunting anti-Semitic tropes, by Russian bots, and by highly visible longtime anti-vaccine campaigners.

Hotez calls for the US federal government to address anti-science aggression and to foster forums that better prepare scientists to act as science communicators. Sadly, an initiative planned to address the latter issue has already been paused, perhaps permanently, by the acting director of the National Institutes of Health, Larry Tabak (*2*). Its advocates fear that the agency has, for political reasons, censored itself because of fears of a right-wing congressional backlash. Thus, the help from the federal government that Hotez hopes for seems unlikely.

The book concludes with some recommendations for how to combat disinformation and anti-science efforts. These include Hotez's plea for the creation of organized support from scientific professional societies for scientists under attack and a proposal for a new entity akin to the Southern Poverty Law Center that would both monitor hateful threats to scientists and offer legal advice and resources.

The book has its weaknesses. For one, the widespread rejection of public health measures during the COVID-19 pandemic is a specific type of anti-scientific behavior that will not necessarily be dispelled by mitigating scientific disinformation or pushing back against right-wing idealogues. A key but overlooked aspect of this fight will be defending the moral, altruistic basis of public health endeavors against a new, virulent form of selfish individualism. Despite some shortcomings, Hotez's warning about the broader implications of Covid denial must be heeded. ■

REFERENCES AND NOTES

1. J. Wallace, P. Goldsmith-Pinkham, J. L. Schwartz, "Excess death rates for Republicans and Democrats during the COVID-19 pandemic," Working Paper 30512 (NBER, 2022); http://www.nber.org/papers/w30512.
2. D. Tahir, "The NIH halts a research project. Is it self-censorship?," CBS News, 5 August 2023; https://www.cbsnews.com/news/nih-halts-research-project-is-it-self-censorship/.

10.1126/science.adi7631

SCIENCE AND SOCIETY

Ableism and assistive technology

Harmful tropes can lead to misguided assumptions when designing devices for people with disabilities

By **Leslie Berntsen**

Against Technoableism: Rethinking Who Needs Improvement
Ashley Shew
Norton, 2023. 160 pp.



Ashley Shew refers to herself as "a hard-of-hearing, chemobrained amputee with Crohn's disease and tinnitus." She is also, by her own admission, a little cranky. One does not have to get very far into her new book, *Against Technoableism*, to understand why. "Ableism"—prejudicial behavior commonly faced by people with disabilities—is a deeply entrenched, but often overlooked, form of structural discrimination. Shew coins the term "technoableism" to describe an extension of this belief: that technological advancements can—and should—be used to eliminate disability.

As Shew notes, the last of the United States' "ugly laws," which penalized people with physical disfigurements for existing in public view, was not repealed until the 1970s. Meanwhile, the Supreme Court case that upheld the government's right to forcibly sterilize an intellectually disabled citizen, *Buck* v. *Bell*, has yet to be formally overturned. At their core, such injustices are informed by a unifying belief: Disabled people—rather than societal attitudes toward disability—are a problem to be solved.

The origins of this belief and its negative consequences for disabled people form the basis of the book's six chapters. Drawing on Shew's personal life and her academic background in disability studies, they touch on a wide array of topics ranging from her day-to-day experiences as a disabled college professor; to her visit to an amputee convention; to the intersection of disability justice, climate change, and space travel.

The technologies that can be used to "fix" disability can take many forms. At the Judge Rotenberg Center in Massachusetts, writes Shew, staff members use electroshock devices to eliminate "undesirable" behaviors in autistic children, a practice that garnered the attention of the United Nations Special Rapporteur on Torture in 2010. (Congress recently granted the US Food and Drug Administration the authority to ban this practice, but it has yet to do so.) Other seemingly more innocuous technologies, such as cochlear implants, which transmit audio data directly to the brains of deaf individuals, are not always embraced by the communities for whom they are intended. As Shew describes in a chapter dedicated to getting her own prosthetic leg, the reality of such interventions is often complicated.

Perhaps one of the book's most important contributions is its discussion of common disability archetypes and their harms. There are the "pitiable freaks" and "inspirational overcomers," who are featured on the local news for being asked to prom by a nondisabled classmate or for running a marathon with their new crowdfunded prosthetic leg. There are the "moochers and fakers" seeking "special" treatment, such as workplace accommodations. There are also the "shameful sinners" who have a disability that they "deserve," such as obesity-related type 2 diabetes.

As Shew notes, and as I can personally attest, being reduced to any of these tropes is frustrating, to say the least, and can easily lead a person with a disability to be cast in an additional role: "the bitter cripple" who describes their experiences in a way that is perceived as off-putting or combative. This final trope is particularly harmful, Shew rightfully argues, because it requires disabled people to advocate for their needs with a smile. But, as Australian comedian Stella Young once pointed out in a TED talk quoted in a later chapter, "No amount of smiling at a flight of stairs has ever made it turn into a ramp" (*1*). By explicitly naming and describing these tropes, Shew helps readers recognize how deeply they are ingrained in public conversations about disability and highlights one of the book's recurring themes: Disabled people too often have our stories told for us, rather than by us.



Tech "fixes" fail to acknowledge the importance of creating inclusive communities.

Disabled people are—and will always be—the experts of our own experiences, but we are not a monolith. In discussing word choice and disabled identity, Shew writes: "Disabled people are a huge and varied group, and we're not all going to agree." This is an important point, and one that readers would do well to keep in mind while considering what it means to oppose technoableism.

Shew makes it clear that she is not opposed to assistive technologies, just to the ways they can function in the service of the aforementioned tropes. Some of the authors whose work she cites hold different opinions, framing the mere existence of these technologies as an affront to the value of disabled lives. Together, these differing viewpoints reinforce the reality that disabled people are not a single, united group.

Against Technoableism contains many lessons, and this is perhaps its most important: Disabled people will not always agree, but we all deserve to be heard. ■

The reviewer is at The Story Collider and based in Los Angeles, CA, USA. Email: leslieb@storycollider.org

REFERENCES AND NOTES

1. S. Young, "I'm not your inspiration, thank you very much," TEDxSydney, April 2014; https://www.ted.com/talks/stella_young_i_m_not_your_inspiration_thank_you_very_much.

10.1126/science.adj4475

PHOTO: VCG VIA GETTY IMAGES



Extreme heat can cause rubber (a polymer) in tires to fail, leading to blowouts.

Edited by **Jennifer Sills**

Editorial Expression of Concern

On 14 October 2022, *Science* published the Research Article "In vivo direct imaging of neuronal activity at high temporospatial resolution" by P. T. Toi *et al.* (*1*). The editors have been made aware that the methods described in the paper are inadequate to allow reproduction of the results and that the results may have been biased by subjective data selection. We are alerting readers to these concerns while the authors provide additional methods and data to address the issues.

H. Holden Thorp
Editor-in-Chief

REFERENCES AND NOTES

1. P. T. Toi *et al.*, *Science* **378**, 160 (2022).

Published online 24 August 2023
10.1126/science.adk4448

Heatwaves hasten polymer degradation and failure

Extreme heatwaves are growing more frequent, and July marked the hottest month ever recorded (*1*). The rising heat increases the risk of premature failure in products made from polymers. Polymer safety specifications must be revisited to ensure their resilience against rising temperatures.

Heat increases molecular mobility and weakens intra- and intermolecular interactions, reducing the structural integrity of plastics, rubbers, fibers, coatings, adhesives, foams, and composites. Increasing temperatures cause these polymers to lose stiffness, strength, creep resistance, barrier properties, electric resistivity, and chemical resistance (*2*). For example, the stiffness of a polyethylene pipe drops by 40% as the temperature rises from 23° to 40°C (*3*).

Heat also accelerates polymer degradation processes, such as oxidation, photodegradation, hydrolysis, and additive migration (*4*, *5*). This leads to faster embrittlement and surface cracking, reducing the lifespan of polymeric products. Polymer degradation intensifies during summer as a result of the higher temperatures, ultraviolet (UV) intensity, humidity, and extended daylight hours (*6*). Heatwaves exacerbate already rapid degradation: Each 10°C temperature rise, within the range seen in extreme heatwaves (*1*), can double the degradation rate (*4*).

Despite engineered safety margins for polymers, heatwaves elevate the risk of premature component or product failure, which could take the form of deformation or warping, brittle fracture, creep rupture, fatigue fracture, environmental stress cracking, delamination due to melting, or discoloration (*7*). Fatigue fracture is responsible for tire blowouts, which can cause severe car accidents (*8*). The melting of polymeric adhesives causes detachment of components such as shoe soles (*9*). Brittle fracture leads to the failure of polyethylene films, which are used extensively in agriculture, bags, and food packaging, after 400 hours at 40°C under UV exposure, and after only 192 hours at 50°C (*10*). Black asphalt and other surfaces have reached more than 80°C in extreme heat (*11*). In addition to posing safety hazards, polymer failures require repair and replacement, leading to economic losses and more plastic waste.

Given escalating heatwaves driven by global warming and El Niño (*12*), revisiting the requirements for polymers to withstand heat is urgent. Governments, polymer producers, scientists, and other stakeholders will have to work together to assess risk, optimize materials, conduct innovative research, ensure regulatory compliance, and monitor and maintain existing products.

Xin-Feng Wei and Mikael S. Hedenqvist*
Division of Polymeric Materials, Department of Fibre and Polymer Technology, KTH Royal Institute of Technology, SE–100 44 Stockholm, Sweden.
*Corresponding author. Email: mikaelhe@kth.se

REFERENCES AND NOTES

1. Copernicus Climate Change Service, "Surface air temperature for July 2023" (2023); https://climate.copernicus.eu/surface-air-temperature-july-2023.
2. U. W. Gedde, M. S. Hedenqvist, M. Hakkarainen, F. Nilsson, O. Das, *Applied Polymer Science* (Springer, 2021).
3. N. Merah, F. Saghir, Z. Khan, A. Bazoune, *Plast. Rubber Compos.* **35**, 226 (2006).
4. B. Singh, N. Sharma, *Polym. Degrad. Stab.* **93**, 561 (2008).
5. A. Stubbins, K. L. Law, S. E. Muñoz, T. S. Bianchi, L. Zhu, *Science* **373**, 51 (2021).
6. G. Grause, M.-F. Chien, C. Inoue, *Polym. Degrad. Stab.* **181**, 109364 (2020).
7. J. Moalli, *Plastics Failure Analysis and Prevention* (William Andrew, 2001).
8. A. Hoyt, "Why do tires blow out more in summer?" *HowStuffWorks* (2022).
9. Y. Dzhanova, "Arizona crossing guard's shoes melt in extreme heat while guiding kids across street," *The Messenger News* (2023).

10. A. Fairbrother *et al.*, *Polym. Degrad. Stab.* **165**, 153 (2019).
11. "How concrete, asphalt and urban heat increase misery of heat waves," *VoA News* (2023).
12. World Meteorological Organization, "World Meteorological Organization declares onset of El Niño conditions" (2023); https://public.wmo.int/en/media/press-release/world-meteorological-organization-declares-onset-of-el-ni%C3%B1o-conditions.

10.1126/science.adj4036

Protest anti-science agenda in Argentina

The candidate who received the most votes in the 13 August Argentinean presidential primary election, Javier Milei, represents the extreme right-wing liberal party, "La Libertad Avanza" (Liberty Advances) (*1*). After winning the election, Milei promised to eliminate the Ministry of Science and the Argentine Research Council (CONICET) and argued that scientists should find work in the private sector instead (*2*).

CONICET, Argentina's main national research and technology institution, employs about 12,000 researchers and a similar number of fellows from diverse fields (*3*). According to the SCImago Institution productivity ratings, CONICET is the second-ranked research institution and the first-ranked governmental institution in Latin America (*4*). It is also ranked among the best 200 research institutions in the world and number 22 among governmental institutions (*4*).

Although the private sector is important in science, the state plays an irreplaceable role (*5*). Nationally funded science is generally not biased toward the private sector's economic interests, aiming instead to improve the well-being of society as a whole. Moreover, the public sector is able to support high-risk basic and applied science over long periods, which also benefits the private sector and society at large (*5*). The importance of government support for science, and the benefits for society, are evident in developed countries, which allocate a large proportion of their gross domestic product to supporting science and technology (*6*) and as a result regularly produce breakthrough science (*5*, *7*).

Argentina is the eighth largest country in the world (*8*), is highly biodiverse (*9*), and has produced three recipients of Nobel Prizes in science (*10*). If Milei, who also denies the existence of climate change and its environmental consequences (*11*), wins the 22 October election, Argentine science, biodiversity, and society may be at risk. The inability of a country to produce its own scientific knowledge limits its capacity to address environmental problems, socioeconomic challenges, and people's well-being (*12*). To cope with scientific denialism, researchers, institutions, and other stakeholders must stand together and protest extreme anti-science agendas. As Argentine recipient of the 1944 Nobel Prize in medicine Bernardo Houssay said, "Science is not expensive; expensive is ignorance" (*10*).

Sergio A. Lambertucci[1]*, Pablo Plaza[1], Julián Padro[1,2], Fernando Valladares[3], Fernando Hiraldo[4], Karina L. Speziale[1]

[1]Grupo de Investigaciones en Biología de la Conservación, Instituto de Investigaciones en Biodiversidad y Medio ambiente (Consejo Nacional de Investigaciones Científicas y Técnicas–Universidad Nacional del Comahue), Quintral 1250, Bariloche, Argentina. [2]Department of Wildlife Sciences, University of Göttingen, 37077 Göttingen, Germany. [3]Museo Nacional de Ciencias Naturales, Consejo Superior de Investigaciones Científicas (CSIC), E-28006 Madrid, Spain. [4]Department of Conservation Biology, Estación Biológica de Doñana, CSIC, Sevilla, Spain.
*Corresponding author.
Email: slambertucci@comahue-conicet.gob.ar

REFERENCES AND NOTES

1. "Far-right outsider takes shock lead in Argentina primary election," *The Guardian* (2023).
2. "Javier Milei confirms he would close CONICET scientific research council," *Buenos Aires Times* (2023).
3. CONICET (2023); https://www.conicet.gov.ar/conicet-descripcion/ [in Spanish].
4. SCImago Ranking Institutions (2023); https://www.scimagoir.com/rankings.php?country=Latin+America&sector=all.
5. A. J. Salter, B. R. Martin, *Res. Pol*. **30**, 509 (2001).
6. Organisation for Economic Cooperation and Development, Research and development (R&D) - Gross domestic spending on R&D - OECD Data (2023); http://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm.
7. A. Gazni, *Technol. Societ*. **62**, 101276 (2020).
8. Worldometer, Largest Countries in the World by Area (2023); https://www.worldometers.info/geography/largest-countries-in-the-world/.
9. J. Caldecott, M. Jenkins, T. Johnson, B. Groombridge, *Biodivers. Conserv*. **5**, 699 (1996).
10. A. Garay, C. Franceschini, S. Valiensi, V. Leske, in *The Practice of Sleep Medicine Around the World: Challenges, Knowledge Gaps, and Unique Needs*, H. P. Attarian, M.-L. M. Coussa-Koniski, A. M. Sabri, Eds. (Bentham Science Publishers, 2023), p. 310.
11. F. R. Molina, "What's going on inside Javier Milei's head?" *EL PAÍS English* (2023).
12. M. Coccia, *Technol. Societ.* **59**, 101124 (2019).

10.1126/science.adk4615

ERRATA

Erratum for the Research Article "Light-induced mobile factors from shoots regulate rhizobium-triggered soybean root nodulation," by T. Wang *et al.*, *Science* **381**, eadj7468 (2023).

Published online 28 July 2023
10.1126/science.adj7468

Erratum for the Research Article "Garnet crystallization does not drive oxidation at arcs" by M. Holycross and E. Cottrell, *Science* **381**, eadj6418 (2023).

Published online 14 July 2023
10.1126/science.adj6418

NEWS FROM
Science
AAAS



RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**



IMMUNOMETABOLISM

Does chitin help metabolic health?

Chitin (β-1,4-poly-*N*-acetylglucosamine) is a highly abundant natural polysaccharide found in arthropods and fungi that can initiate type 2 (allergic) immune responses. Kim *et al.* report in mice that the consumption of chitin triggers gastric distension, downstream neuropeptide release, and type 2 cytokine production by tuft cells and type 2 innate lymphoid cells. This in turn drives gastrointestinal remodeling and the generation of acidic mammalian chitinase by chief cells, which is needed for chitin digestion. The addition of dietary chitin improved metabolic readouts in mice fed a high-fat diet, possibly because activated chief cells produce other digestive enzymes, including lipase. This mammalian adaptation to chitin may therefore serve as a potential therapeutic target for metabolic diseases such as obesity. —STS

Science, add5649, this issue p. 1092

Chitin, a structural polysaccharide found in some foods such as mushrooms, stimulates the release of digestive enzymes in the stomach through a type 2 immune circuit.

PHOTO: NATASHA BREEN/REDA&CO/UNIVERSAL IMAGES GROUP/GETTY IMAGES

BROWN DWARFS

A radiation belt around a brown dwarf

In the Solar System, planets with strong magnetic fields trap energetic plasma above their equators, forming a radiation belt. Theory has predicted that this process should also occur for brown dwarfs, objects with masses between those of planets and stars. Climent *et al.* have observed a nearby brown dwarf using radio interferometry. They identified spatially resolved emission that varied as the brown dwarf rotated. Comparing these observations with models, the authors believe that a radiation belt is present around the brown dwarf and constrains its properties. The results indicate similarities between the magnetospheres of brown dwarfs and gas giant planets. —KTS

Science, adg6635, this issue p. 1120

ORGANOMETALLICS

Copper surprises

Copper-mediated couplings are among the oldest reactions in organic chemistry. Nonetheless, their mechanisms still are not as predictably well understood as those of the palladium catalysts that came later. Part of the difficulty is tracking the one-electron increments through which copper typically reacts, shuffling between the +1 and +2 oxidation states. Two studies now uncover two-electron processes at play instead. Delaney *et al*. found that ligand oxidation enables oxidative addition of an aryl halide to Cu(II), avoiding the need for a less stable Cu(I) precursor. Luo *et al*. observed halonitrile addition to Cu(I) complexes to produce isolable Cu(III). Both results point to future innovative catalyst optimization strategies. —JSY

Science, adg9232, adi9226, this issue p. 1072, p. 1079

MEMBRANES

Tunable high-flux inorganic membranes

Polymer membranes are commonly used for a range of molecular separations, but they do not work well under elevated temperatures or with harsh solvents. Sengupta *et al*. describe a method to make inorganic membranes that is analogous to interfacial polymerization. Chemical vapor deposition is used to fabricate carbon-doped metal oxide interfacial nanofilms on ceramic substrates with pore sizes that can be tuned by subsequent sintering. The membranes, although thicker

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

GREEN NUDGES

Reducing waste for food orders

China's high demand for online food delivery resulted in an increase in the use of disposable, single-use cutlery. Disposable cutlery increases plastic pollution, and paper napkins and wooden chopsticks contribute to environmental degradation that endangers wildlife and marine species and compromises human health. Informed by the literature on "green nudges," which are prompts to promote environmentally friendly behaviors, He *et al*. collaborated with Alibaba to use its mobile food delivery platform, Eleme, in a longitudinal field study across China. Forgoing cutlery became the app's default option, was made salient to users in a pop-up window, and was rewarded with "green points" redeemable for planting trees in China's deserts. The results suggest that incorporating green nudges is a feasible and effective policy tool because it greatly increased orders forgoing single-use cutlery without harming restaurants' business performance. —EEU

Science, add9884, this issue p. 1064

CANCER

Stressing out escapees

Most cancers depend on a balanced proteostasis network to maintain oncogenic growth, and therapeutic insults often disrupt proteostasis. However, how residual drug-tolerant cells overcome imbalances in the proteostasis network to survive targeted therapy is poorly understood. Lv *et al*. discovered a mechanism that rewires the proteostasis network to escape oncogenic KRAS addiction and promote resistance to KRAS inhibitors (KRASi). Inactivation of mutant KRAS, a key oncogenic driver for many human cancers, shuts down major proteostasis regulatory pathways, causing severe protein aggregations. However, multiple resistance mechanisms converge to selectively reactivate an ancient stress sensor, IRE1α, to reestablish proteostasis in the KRASi-resistant cells. Targeting IRE1α could collapse the rewired proteostasis network and overcome resistance to KRASi therapy. —SMH and PNK

Science, abn4180, this issue p. 1065

IMMUNOLOGY

All together now!

Kupffer cells (KCs) are specialized macrophages in the liver sinusoids that filter bacteria in the bloodstream. Liver fibrosis and cirrhosis, a common pathology of chronic liver disease, leads to the redistribution of blood flow from the sinusoids to collateral vessels. Using a mouse model of hepatic fibrosis, Peiseler *et al*. report that remodeling of the liver causes KCs to lose contact with surrounding cells and lose their distinct cellular identity (see the Perspective by Louwe and Guilliams). The liver still maintains its bacterial filtration function because the presence of a microbiome helps to recruit monocytes to large intrahepatic vessels, where they fuse into large cell aggregates that exhibit a KC-like phenotype. These KC-like syncytia show enhanced bacterial capture capacity and are present in cirrhotic livers from patients with various chronic liver diseases. —STS

Science, abq5202, this issue p. 1066; see also adj9725, p. 1050

VOLCANOLOGY

Destructive density current

Volcanic eruptions can produce dangerous, destructive, fast-moving flows of ash and rock. The 2022 Hunga Tonga–Hunga Ha'apai eruption in Tonga also produced massive amounts of ash and rock, but with the destructive flow occurring underwater. Clare *et al.* found that the volcano produced a massive, fast-moving underwater debris flow that traveled more than 100 kilometers (see the Perspective by Williams and Rowley). The flow reshaped the seafloor and destroyed both international and domestic telecommunications cables. The speed and size of the flow were not expected, and such flows present a risk scenario for undersea telecommunications cables that has not been previously recognized. —BG

Science, adi3038, this issue p. 1085; see also adk0181, p. 1046

NEUROSCIENCE

Lifelong cerebellar remodeling

Containing nearly half of the neurons in the brain, the cerebellum is one of its most neuron-dense areas. Tan *et al*. performed transcriptome and chromatin accessibility analysis at different developmental stages in human and mouse cerebellar granule cells (see the Perspective by Akbarian and Won). The authors describe the main differences between human and mouse granule cells, but also point out that both show continuous modulation of genome architecture throughout life. These results provide valuable insights into granule cell changes during development and aging. —MMa

Science, adh3253, this issue p. 1112; see also adk0961, p. 1049

DENDRITIC CELLS

Steering clear

Dendritic cells (DCs) migrate over long distances to shuttle antigen from peripheral tissues to lymph nodes. Immune cell trafficking is mediated by chemotactic gradients, but whether DCs can modulate these gradients is unknown. Using a combination of live-cell imaging and mathematical modeling, Alanko *et al*. identified a dual role for the G protein–coupled receptor CCR7 in controlling DC migration. In addition to sensing the chemokine ligand CCL19 through CCR7, DCs were able to modulate local chemokine concentration through endocytosis of CCR7, "sinking" CCL19 from the environment. This self-shaping of chemokine gradients facilitated the accurate migration of DCs, and these gradients could also be sensed by T cells. These findings demonstrate that CCR7 can function as both a sensor and a sink for CCL19, facilitating collective leukocyte migration. —HMI

Sci. Immunol. (2023) 10.1126/sciimmunol.adc9584

than others currently in use, still show ultrafast permeance of pure solvents. Furthermore, they can operate in harsh solvents at temperatures of up to 140°C. —MSL

Science, adh2404, this issue p. 1098

COMMUNITY ECOLOGY

Selection for small size

Large organisms may be particularly susceptible to impacts from human activities, including hunting and harvesting, as well as the stress of climate change. Martins *et al*. analyzed body-size trends in plant and animal communities since 1960 from the BioTIME database and separated the effects of changes within species from those driven by species composition. They found that trends varied across communities, with marine fish more consistently shifted toward smaller body size. Mean body size changed within populations, but community-level trends were more affected by changes in the abundance of small- and large-bodied species. Biomass generally stayed stable over time, suggesting that more small individuals were present despite the loss of larger ones. —BEL

Science, adg6006, this issue p. 1067

AUTOIMMUNITY

The source of autoimmune vasculitis

Giant cell arteritis (GCA) is an autoimmune disease of the medium and large arteries that affects older adults and is often refractory to standard immunosuppression. Although clusters of T cells occupy the vessel walls in patients with GCA, the source of these cells and their importance in driving disease is unclear. Using both human aortic tissues and serial transplantation experiments in mice, Sato *et al*. found that a population of stem-like $CD4^+$ T cells residing in tertiary lymphoid structures replenishes pathogenic effector cells independently of secondary lymphoid organs. Targeting such stem-like $CD4^+$ T cells may represent a therapeutic opportunity to halt the chronic progression of autoimmune vasculitis. —CM

Sci. Transl. Med. (2023)
10.1126/scitranslmed.adh0380

BIOMEDICINE

Early detection of kidney transplant failure

Organ transplantation can considerably extend the recipient's life, but organ failure can occur at any time, especially if the host rejects the organ. Madhvapathy *et al*. developed a bioelectronics-based approach to track acute kidney organ transplant failure by attaching a wireless, soft electronics interface to transplanted kidneys that allows real-time continuous monitoring of organ temperature and thermal conductivity (see the Perspective by Zaidan and Lakkis). Using rat models, the authors demonstrated early detection of transplant rejection by tracking changes in normal cycles or an elevation in organ temperature. —MSL

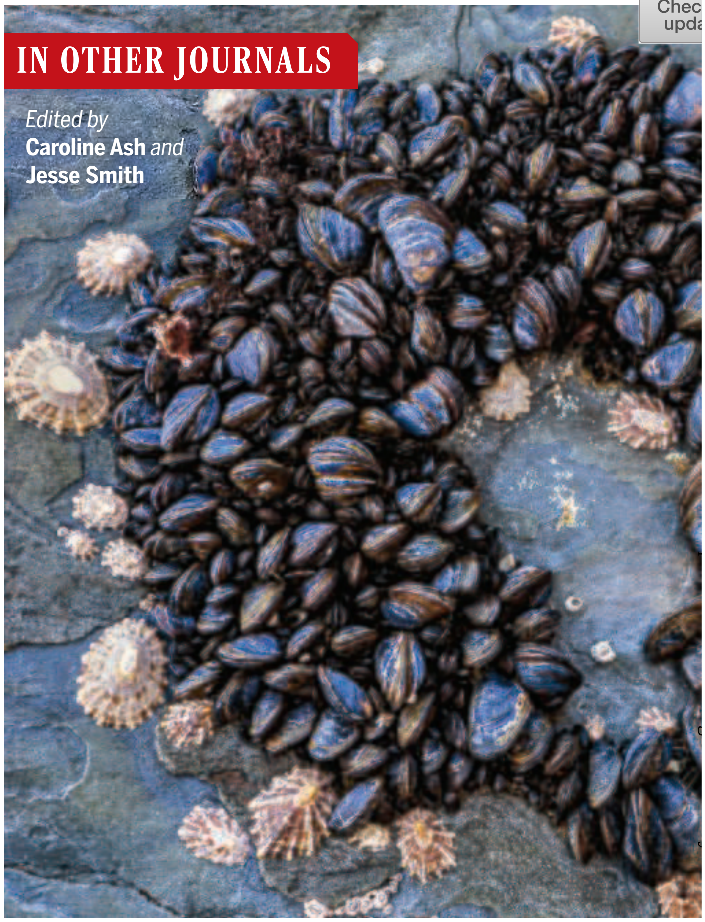
Science, adh7726 this issue p. 1105;
see also adj9517, p. 1048

MICROBIOLOGY

Lactylation in bacteria on a sugar high

High sucrose triggers biofilm formation and the production of tooth-damaging lactate by the oral bacterium *Streptococcus mutans*. Li *et al.* identified and characterized lysine residues in *S. mutans* that were lactylated (posttranslationally modified with lactate). High sucrose generally elicited a global increase in lysine lactylation, particularly in proteins involved in metabolism and protein synthesis. In vitro and in vivo experiments identified an acetyltransferase that limited biofilm formation during high-sucrose conditions, potentially by lactylating an RNA polymerase subunit. —AMV

Sci. Signal. (2023)
10.1126/scisignal.adg1849



IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**

PSYCHOLOGY

My beautiful (future) friend

People tend to evaluate themselves and others who are close to them more positively than they objectively are. When people select others to befriend, they tend to anticipate someone who looks, smells, and sounds better than they are. Kononov *et al.* were interested in understanding how we use these physical parameters to evaluate strangers. The results, obtained from 401 participants from the United States, showed that the amount of time participants expected to spend with the stranger in the future was correlated with the magnitude of enhanced evaluation. The desire for a pleasant experience seems to result in overly favorable perceptions of strangers. —MMa

Pers. Soc. Psychol. Bull. (2023)
10.1177/01461672231180150

REPRODUCTION

Building the ovarian reserve

In mammals, the entire lifetime supply of ovarian follicles is established before birth. In some cases, follicles may become depleted prematurely, leading to infertility. By investigating the functions of orphan nuclear receptors, Hughes *et al*. found that one of them, steroidogenic factor 1 (SF-1), contributes to the prenatal development and subsequent maintenance of the ovarian reserve in multiple species. The full mechanism of action of SF-1 is not yet known, but it appears to have several beneficial effects on oocytes, including the formation of appropriate ovarian follicle structure and connections, as well as reductions in the rates of lysosomal degradation and autophagic death. —YN

Proc. Natl. Acad. Sci. U.S.A. (2023)
10.1073/pnas.2220849120



MATERIALS SCIENCE

Efficient selective neodymium recovery

Neodymium is a key rare earth element and an essential component for some strong magnets and lasers. Recovery of neodymium from recycled electronics or industrial waste is possible using adsorbents in aqueous media, but recovery rates are low. Yeh *et al.* drew inspiration from mussels by fabricating cellulose-based "hairy cellulose" nanocrystals that can adhere to silica gel microparticles with the help of dopamine and selectively remove neodymium ions even with other ions present. The decorated microparticles show excellent selective adsorption of neodymium ions, are easier to recover from solution, and can be reused for multiple recovery cycles. —MSL

ACS Appl. Mater. Interfaces (2023)
10.1021/acsami.3c04512

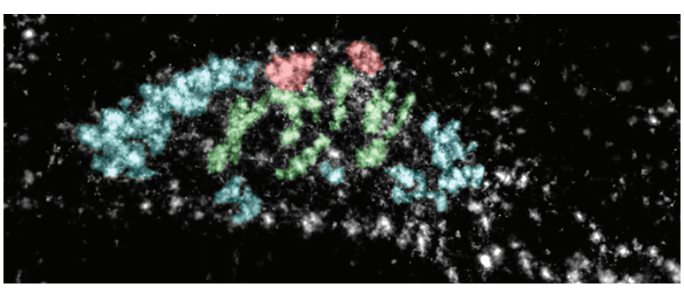
Synthetic "hairy cellulose nanocrystals" inspired by adhesive proteins found in mussels can be used to recover neodymium from waste streams.

CELL BIOLOGY

Presynaptic actin organization

Actin, together with myosin, is an abundant contractile protein in muscle fibers and is required for cell movement. Actin is highly expressed in the presynaptic regions of neurons. Bingham *et al*. specifically visualized presynaptic regions of cultured rat neurons using superresolution imaging. A subpopulation of presynapses were enriched in actin, which correlated with higher synaptic vesicle cycling. What is known as single-molecule localization microscopy of presynapses revealed three distinct types of actin nanostructures. At the active zone, actin formed into a mesh-like structure. Between the active zone and the synaptic vesicle reserve pool, deeper within the presynapse, actin formed into rail-like structures. Finally, the whole presynaptic compartment was enclosed in a corral-like structure of actin. These elaborate presynaptic actin nanostructures may be important for efficient neurotransmission. —SMH

J. Cell Biol. (2023)
10.1083/jcb.202208110



DNA-PAINT image of actin in the presynapse of a rat neuron showing the active zone mesh (red), actin rails (green), and the enclosing corral-like structure (blue)

PHOTO: D. BINGHAM *ET AL*., *JOURNAL OF CELL BIOLOGY* 10.1083/JCB.202208110 (2023)

GENOME EDITING

Miniature genome editors

CRISPR-Cas9 systems have been developed into powerful genome-editing tools that have had revolutionary effects in medicine and agriculture. Cas9 is a large, 160-kilodalton protein, and smaller nucleases would increase delivery efficiency. Recent successes have come from tracing the evolutionary ancestors of Cas9 and related Cas12 proteins. Xiang *et al.* have found one endonuclease ancestor, TnpB, and identified two small motifs, ISAam1 and ISYmu1, that have robust editing activity across dozens of genomic loci in human cells. These are 369 and 382 amino acids long, respectively, about one-third of the length of the most widely used Cas9. These evolutionary ancient editors have the potential to become effective genome-editing tools. —DJ

Nat. Biotechnol. (2023)
10.1038/s41587-023-01857-x

GALAXIES

Simulated galaxies in the early Universe

Previous galaxy simulations correctly matched the number of observed galaxies from the local Universe back to redshifts (z) of about 8 (about 600 million years after the Big Bang). Recent observations have detected more galaxies at $z \geq 8$ than were expected. Kannan *et al.* used the MillenniumTNG hydrodynamic simulations to predict the number of galaxies expected at $z = 7$ to 15. Their simulations agree with observations back to $z \sim 10$, and to $z \sim 12$ if there is low dust abundance. However, at even higher redshifts (less than 400 million years after the Big Bang), the simulations underpredict the observed number of galaxies. The authors suggest that this could be due to more efficient star formation at such early times. —KTS

Mon. Not. R. Astron. Soc. (2023)
10.1093/mnras/stac3743

ORGANIC CHEMISTRY

Halides reconfigure a ruthenium catalyst

Asymmetric catalysis often relies on the configuration of a chiral ligand coordinated to a metal center. Shezaf *et al.* report an unusual instance in which the configuration at the metal itself, in addition to the chiral ligand, plays a decisive role in the reaction. Specifically, chloride coordinates opposite a carbon monoxide ligand on a ruthenium center with a chiral phosphine, whereas iodide coordinates perpendicular to it. As a result, the catalyst changes its regioselectivity in coupling an alcohol to isoprene, favoring a bond to a secondary carbon center with chlorine and a tertiary carbon center with iodine. —JSY

J. Am. Chem. Soc. (2023)
10.1021/jacs.3c06734

0908ISIO_17769502.indd 1063 9/1/23 4:16 PM

RESEARCH ARTICLE SUMMARY

CANCER

Modulation of the proteostasis network promotes tumor resistance to oncogenic KRAS inhibitors

Xiangdong Lv†, Xuan Lu†, Jin Cao†, Qin Luo, Yao Ding, Fanglue Peng, Apar Pataer, Dong Lu, Dong Han, Eric Malmberg, Doug W. Chan, Xiaoran Wang, Sara R. Savage, Sufeng Mao, Jingjing Yu, Fei Peng, Liang Yan, Huan Meng, Laure Maneix, Yumin Han, Yiwen Chen, Wantong Yao, Eric C. Chang, Andre Catic, Xia Lin, George Miles, Pengxiang Huang, Zheng Sun, Bryan Burt, Huamin Wang, Jin Wang, Qizhi Cathy Yao, Bing Zhang, Jack A. Roth, Bert W. O'Malley, Matthew J. Ellis, Mothaffar F. Rimawi, Haoqiang Ying*, Xi Chen*

INTRODUCTION: KRAS is one of the most frequently mutated genes in human cancer. Despite advances in the development of inhibitors that directly target mutant KRAS and the approval of KRASG12C inhibitors sotorasib and adagrasib for the treatment of KRASG12C-mutant non–small cell lung cancer (NSCLC) patients, multiple lines of clinical and preclinical evidence demonstrate that adaptive resistance to KRAS inhibitors (KRASi) is rapid and almost inevitable. The heterogeneous resistance mechanisms in patients and dose-limiting toxicity associated with targeting multiple KRASi resistance pathways—such as receptor tyrosine kinases (RTKs), extracellular signal–regulated kinase (ERK), and AKT–remain a major barrier to progress.

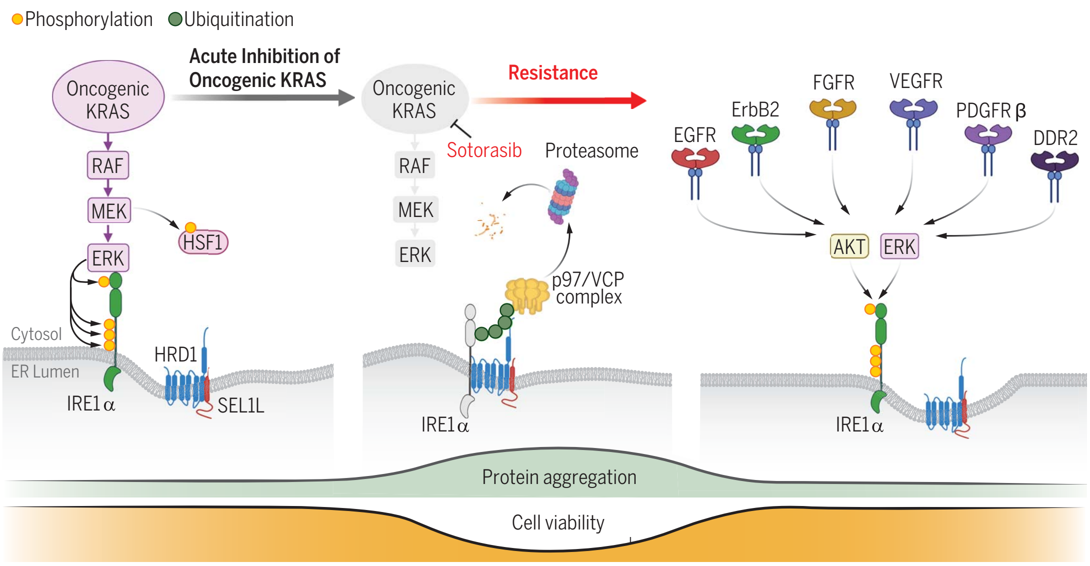
RATIONALE: Most cancers require a balanced protein homeostasis (proteostasis) network to maintain oncogenic growth. Therapeutic insults often disrupt proteostasis and induce proteotoxic stresses. Residual drug-tolerant cells must overcome imbalances in the proteostasis network to maintain survival. How a proteostasis network is orchestrated by driver oncogenes and the proteostasis reprogramming mechanisms that bypass oncogene addiction and allow for acquired resistance to targeted therapies remain largely unknown. In this study, we investigated the regulation of proteostasis by oncogenic KRAS and the rewiring of proteostasis network underlying the acquired resistance to KRAS inhibition.

RESULTS: We show that oncogenic KRAS is critical for protein quality control in cancer cells. Genetic or pharmacological inhibition of oncogenic KRAS rapidly inactivated both cytosolic and endoplasmic reticulum (ER) protein quality control machinery, two essential components of the proteostasis network, through inhibition of the master regulators heat shock factor 1 (HSF1) and inositol-requiring enzyme 1α (IRE1α). However, residue cancer cells that survive KRASi directly reactivated IRE1α through an ER stress–independent phosphorylation mechanism that reestablished proteostasis and sustained acquired resistance to KRAS inhibition. We identified four oncogenic signaling–regulated phosphorylation in IRE1α (Ser^{525}, Ser^{529}, Ser^{549}, and Thr^{973}) that are distinct from IRE1α autophosphorylation sites but are required for enhanced protein stability. The phosphorylation of IRE1α at these sites prevents IRE1α binding with the SEL1L/HRD1 E3 ligase complex, thus impairing the ubiquitination-dependent degradation of IRE1α and stabilizing the protein. These sites are the convergence points of multiple resistance mechanisms in KRASi-resistant tumors. RTK-mediated reactivation of ERK and hyperactivation of AKT sustained the unconventional phosphorylation of IRE1α in the KRASi-resistant tumors, which consequently restored its protein stability and reestablished proteostasis. Genetic or pharmacological suppression of IRE1α collapsed the rewired proteostasis network and overcame resistance to KRAS–MAPK (mitogen-activated protein kinase) inhibitors.

CONCLUSION: This study reveals the direct cross-talk between oncogenic signaling and the protein quality control machinery and uncovers the mechanisms that account for the proteostasis rewiring in response to KRAS inhibition. Multiple resistance mechanisms converge on IRE1α through ER stress–independent phosphorylation to restore proteostasis and promote KRASi-resistant tumor growth. Targeting this key convergence point represents an effective therapeutic strategy to overcome KRASi resistance. ■

The list of author affiliations appears in the full article online.
*Corresponding author. Email: xi.chen@bcm.edu (X.C.); hying@mdanderson.org (H.Y.)
†These authors contributed equally to this work.
Cite this article as X. Lv *et al.*, *Science* **381**, eabn4180 (2023). DOI: 10.1126/science.abn4180

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abn4180



Proteostasis reprogramming upon KRAS inhibition. Inhibition of oncogenic KRAS inactivates both cytosolic and ER protein quality control machinery by inhibiting HSF1 and IRE1α. Residual cells that survive KRASi directly restore IRE1α phosphorylation through receptor tyrosine kinase–mediated reactivation of ERK and hyperactivation of AKT, preventing IRE1α from SEL1L/HRD1–mediated ubiquitination and degradation. Multiple heterogeneous resistance pathways converge on IRE1α to reestablish proteostasis and promote resistance to KRASi.

RESEARCH ARTICLE

CANCER

Modulation of the proteostasis network promotes tumor resistance to oncogenic KRAS inhibitors

Xiangdong Lv[1,2,3]†, Xuan Lu[1,2,3]†, Jin Cao[1,2,3]†, Qin Luo[1,2,3], Yao Ding[1,2,3], Fanglue Peng[1,2,3], Apar Pataer[4], Dong Lu[1,5,6], Dong Han[1,2,3], Eric Malmberg[1,2,3], Doug W. Chan[1,2,3], Xiaoran Wang[1,2,3], Sara R. Savage[2,3,7], Sufeng Mao[1,2,3], Jingjing Yu[1,2,3], Fei Peng[1,8], Liang Yan[9], Huan Meng[1], Laure Maneix[1,10], Yumin Han[1,2,3], Yiwen Chen[11], Wantong Yao[12], Eric C. Chang[1,2,3], Andre Catic[1,10], Xia Lin[13], George Miles[2,3,7], Pengxiang Huang[1], Zheng Sun[1,8], Bryan Burt[14], Huamin Wang[15], Jin Wang[1,5,6], Qizhi Cathy Yao[13], Bing Zhang[2,3,7], Jack A. Roth[4], Bert W. O'Malley[1], Matthew J. Ellis[2,3,16], Mothaffar F. Rimawi[2,3], Haoqiang Ying[9]*, Xi Chen[1,2,3]*

Despite substantial advances in targeting mutant KRAS, tumor resistance to KRAS inhibitors (KRASi) remains a major barrier to progress. Here, we report proteostasis reprogramming as a key convergence point of multiple KRASi-resistance mechanisms. Inactivation of oncogenic KRAS down-regulated both the heat shock response and the inositol-requiring enzyme 1α (IRE1α) branch of the unfolded protein response, causing severe proteostasis disturbances. However, IRE1α was selectively reactivated in an ER stress–independent manner in acquired KRASi-resistant tumors, restoring proteostasis. Oncogenic KRAS promoted IRE1α protein stability through extracellular signal–regulated kinase (ERK)–dependent phosphorylation of IRE1α, leading to IRE1α disassociation from 3-hydroxy-3-methylglutaryl reductase degradation (HRD1) E3-ligase. In KRASi-resistant tumors, both reactivated ERK and hyperactivated AKT restored IRE1α phosphorylation and stability. Suppression of IRE1α overcame resistance to KRASi. This study reveals a druggable mechanism that leads to proteostasis reprogramming and facilitates KRASi resistance.

KRAS is one of the most frequently mutated genes in human cancer, especially in pancreatic ductal adenocarcinoma (PDAC), non–small cell lung cancer (NSCLC), and colorectal carcinoma (CRC) (*1*–*5*). Oncogenic KRAS engages multiple effector pathways to drive tumorigenesis, notably the RAF/MEK/ERK (MAPK) and phosphatidylinositol 3-kinase (PI3K)/AKT pathways (*6*–*10*). Small molecules that directly target the *KRAS* G12C mutation (in which glycine at position 12 is replaced by cysteine), including sotorasib and adagrasib (*11*–*13*), have shown encouraging therapeutic efficacies in clinical trials (*14*–*16*). The US Food and Drug Administration (FDA) granted accelerated approval for sotorasib and adagrasib to treat patients with $KRAS^{G12C}$-mutant NSCLC. However, resistance to these KRAS inhibitors (KRASi) is almost inevitable, resulting from the activation of compensatory pathways—such as epidermal growth factor (EGF), fibroblast growth factor (FGF), aurora kinase A (AURKA), or SOS1—or the acquisition of new mutations, such as *KRAS*, *NRAS*, *BRAF*, *EGFR*, or *FGFR2* (*15*–*35*). Although inhibitors that target *KRAS* mutations other than G12C are in clinical trials (*36*), similar bypass of *KRAS* dependence has been demonstrated in preclinical studies that used genetically engineered mouse models (GEMMs) of lung and pancreatic cancer (*37*–*39*). Because of these clinical challenges, understanding the mechanisms that mediate the resistance to KRASi is imperative to develop more effective therapies to prevent the recurrence of *KRAS*-driven cancers.

The ability to overcome an imbalanced protein homeostasis or "proteostasis" network is instrumental to maintain cancer cell survival and circumvent various insults that impair the quality of protein synthesis and folding (*40*, *41*). Cells use compartment-specific stress sensors to monitor and maintain a high-fidelity proteome. Cytosolic proteins are monitored by the heat shock response (HSR) (*42*), whereas transmembrane and secreted proteins are monitored in the endoplasmic reticulum (ER) by the unfolded protein response (UPR) (*43*–*45*). When the HSR is triggered by stress stimuli, heat shock factors (HSFs) become activated and transactivate genes that encode chaperones and other factors of the proteostasis network (*46*). HSF1 is a master regulator of the proteotoxic stress response and has been implicated in mediating proteome fidelity in cancer cells (*47*). The UPR is a three-branched stress response that is activated upon disruption of ER homeostasis. The UPR is mediated by the ER transmembrane sensors inositol-requiring enzyme 1α (IRE1α), activated transcription factor 6 (ATF6), and protein kinase RNA-like ER kinase (PERK) (*48*). IRE1α is the most ancient and conserved member of the mammalian UPR sensory triad (*49*, *50*). Under ER stress conditions, IRE1α undergoes oligomerization and trans-autophosphorylation to activate its ribonuclease (RNase) domain. This results in excision of 26 nucleotides from unspliced *XBP1* (*XBP1u*) mRNA, and a frame shift mutation to produce the mature, spliced *XBP1* (*XBP1s*), which encodes a potent transcriptional activator (fig. S1A) (*51*, *52*). Mutant RAS was traditionally viewed as an inducer of general UPR through stressing ER during oncogenic transformation of nonmalignant cells (*53*). Genetic screens revealed the synthetic lethal interaction between mutant RAS and IRE1α in yeast (*54*). Yet it remains unclear how the proteostasis network is orchestrated by oncogenic KRAS, or how the proteostasis reprogramming mechanisms occur that bypass KRAS addiction and allow for acquired resistance to KRASi.

Here, we report that proteostasis is dynamically altered upon oncogenic KRAS inhibition. We identified the IRE1α-mediated reprogramming of proteostasis as an essential mechanism that facilitates resistance to KRAS-MAPK inhibition. We also elucidated the biochemical basis for the ER stress–independent posttranslational modification and regulation of IRE1α through both oncogenic KRAS signaling and KRASi resistance signaling, which serves as a therapeutic vulnerability that can be targeted to overcome resistance to KRAS-MAPK inhibition.

Oncogenic KRAS inactivation reprograms proteostasis

To understand the impacts of oncogenic KRAS on proteostasis, we used primary mouse PDAC cells derived from our previously generated,

[1]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA. [2]Lester and Sue Smith Breast Center, Baylor College of Medicine, Houston, TX 77030, USA. [3]Dan L. Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA. [4]Department of Thoracic and Cardiovascular Surgery, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [5]Department of Pharmacology and Chemical Biology, Baylor College of Medicine, Houston, TX 77030, USA. [6]Center for Drug Discovery, Baylor College of Medicine, Houston, TX 77030, USA. [7]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. [8]Department of Medicine, Division of Diabetes, Endocrinology and Metabolism, Baylor College of Medicine, Houston, TX 77030, USA. [9]Department of Molecular and Cellular Oncology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [10]Huffington Center on Aging, Baylor College of Medicine, USA. [11]Department of Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [12]Department of Translational Molecular Pathology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [13]Division of Surgical Oncology, Michael E. DeBakey Department of Surgery, Baylor College of Medicine, Houston, TX 77030, USA. [14]Division of Thoracic Surgery, Michael E. DeBakey Department of Surgery, Baylor College of Medicine, Houston, TX 77030, USA. [15]Department of Pathology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [16]Early Oncology, Oncology R&D, AstraZeneca, Gaithersburg, MD, USA.
***Corresponding author. Email: xi.chen@bcm.edu (X.C.); hying@mdanderson.org (H.Y.)**
†These authors contributed equally to this work.

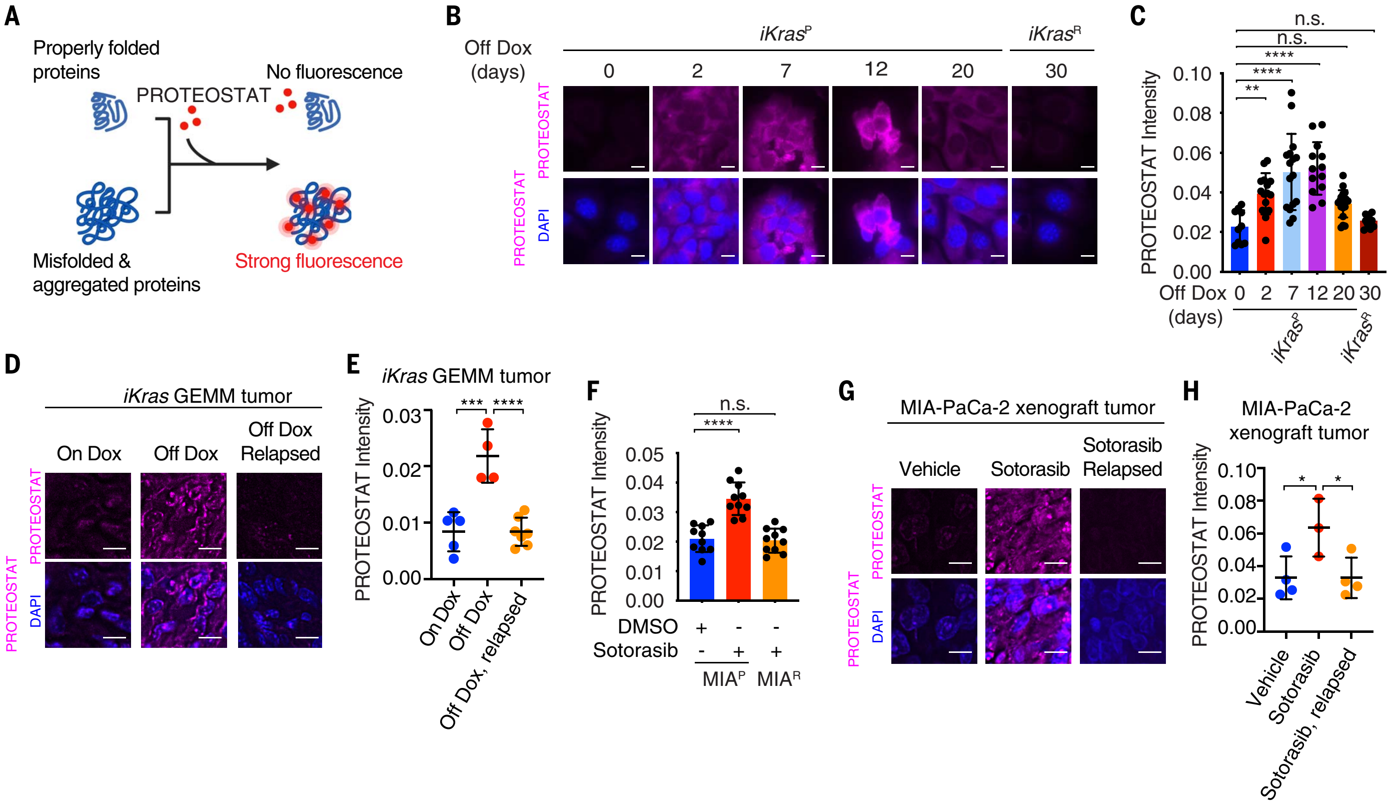


Fig. 1. Oncogenic KRAS inactivation reprograms proteostasis. (**A**) Schematic illustration of labeling and detection of misfolded and aggregated proteins with PROTEOSTAT dye. Upon intercalation into the cross-β spine typically found in misfolded and aggregated proteins, PROTEOSTAT dye emits strong fluorescence. (**B**) Representative images and (**C**) quantification of PROTEOSTAT (magenta) and 4′,6-diamidino-2-phenylindole (DAPI) (blue) staining in *iKras*^{P} cells at different time points after *Kras*^{G12D} inactivation through Dox-withdrawal (Off Dox) until the cells acquired resistance to *Kras*^{G12D} inactivation (*iKras*^{R} cell). (**D**) Representative images and (**E**) quantification of PROTEOSTAT (magenta) staining in spontaneous tumors from the Dox-inducible, *Kras*^{G12D}-driven PDAC mouse model (*iKras* GEMM) treated with doxycycline (Dox, 2 g/liter, *n* = 5 mice), Dox withdrawal for 3 days (*n* = 4 mice) or relapsed after 30 weeks of Dox-withdrawal (*n* = 7 mice). (**F**) Quantification of PROTEOSTAT intensity in parental MIA-PaCa-2 (MIA^P) cells treated with dimethyl sulfoxide (DMSO) or 30 nM sotorasib for 2 days or in sotorasib-resistant MIA-PaCa-2 (MIA^R) cells treated with 30 nM sotorasib. MIA^R cells were generated in vitro by means of continued sotorasib treatment until the cells acquired resistance. (**G**) Representative images and (**H**) quantification of PROTEOSTAT (magenta) and DAPI (blue) staining in MIA-PaCa-2 xenograft tumors treated with vehicle (*n* = 4 mice), sotorasib (30mg/kg for 1 day, *n* = 3 mice), or relapsed after 9 weeks of sotorasib treatment (30mg/kg, *n* = 4 mice). Data represent average fluorescence intensity of PROTEOSTAT per cell from each image [(C) and (F)] or tumor [(E) and (H)] and are presented as mean ± SD from *n* ≥ 10 images. Scale bar, 20 μm. Ordinary one-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test in (C), (E), (F), and (H) was used to calculate *P* values. n.s., not significant, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$.

doxycycline (Dox)–inducible, *Kras*^{G12D}-driven PDAC mouse model (*55*), which are hereafter designated as *iKras* cells. Dox withdrawal turned off *Kras*^{G12D} expression in *iKras* cells (fig. S1B), resulting in the inactivation of downstream MAPK signaling (fig. S1C), decreased 5-bromodeoxyuridine (BrdU) incorporation (fig. S1D), and increased cell apoptosis (fig. S1E). We monitored protein aggregation upon *Kras*^{G12D} inactivation using PROTEOSTAT, a molecular rotor dye that specifically intercalates into the cross-β spines of the quaternary protein structures found in misfolded and aggregated proteins (Fig. 1A). As a positive control, the treatment of *iKras* cells with the proteasome inhibitor MG132 significantly induced PROTEOSTAT signal (fig. S1F). The inactivation of *Kras*^{G12D} from Dox withdrawal induced protein aggregation (Fig. 1, B and C). However, 30 days after Dox withdrawal, *iKras* cells displayed restored proteostasis (Fig. 1, B and C) and resumed cell growth (fig. S1, D, E, and G), at levels comparable with those observed for parental cells (*iKras*^{P}). These cells resistant to *Kras*^{G12D} inactivation are hereafter designated as *iKras*^{R} cells. Continuous *Kras*^{G12D} inactivation was required to maintain the resistance phenotypes because reactivation of *Kras*^{G12D} for extended periods (12 days) partially reversed the resistance of *iKras*^{R} cells to Dox withdrawal (fig. S1, H to K). Dox administration had no impact on cell growth or proteostasis in *LSL-Kras*^{G12D} cells that lack Dox-inducible *Kras*^{G12D} expression (fig. S1, L to P). In the *iKras* cells, staining with Congo red (CR) or thioflavin T (ThT), which recognize misfolded protein aggregates (*56*), or detection of lysine 48 (K48)–linked polyubiquitin levels, which tag misfolded or aggregated proteins for proteasome-mediated degradation (*57*), independently confirmed the restoration of proteostasis upon acquired resistance to *Kras*^{G12D} inactivation (fig. S1, Q to S).

We also used the *iKras* GEMM to examine in vivo proteostasis reprogramming (*55*). *Kras*^{G12D} inactivation through Dox withdrawal resulted in rapid tumor regression, but 70% of tumors relapsed after 9 to 47 weeks (*37*). PROTEOSTAT staining of *iKras* GEMM tumors at different stages of *Kras*^{G12D} inactivation revealed increased protein aggregation in regressing parental tumors and restored proteostasis in relapsed tumors (Fig. 1, D and E, and fig. S2A), which grow independently of *Kras*^{G12D} expression.

To further investigate the impacts of oncogenic KRAS inhibition on proteostasis, we used a KRAS^{G12C} inhibitor, sotorasib (*12*), to treat MIA-PaCa-2 human PDAC cells and H358 human NSCLC cells harboring the *KRAS*^{G12C}

mutation. Sotorasib effectively suppressed the activation of MEK/ERK (fig. S2, B and C), leading to cell growth inhibition in both MIA-PaCa-2 and H358 cell lines (fig. S2, D and E). Both MIA-PaCa-2 and H358 cells also displayed increased PROTEOSTAT staining and K48-linked polyubiquitination levels in response to sotorasib treatment, suggesting enhanced protein aggregation (Fig. 1F and fig. S2, F to H). However, after MIA-PaCa-2 and H358 cells gained resistance to sotorasib (fig. S2, D and E), they also exhibited restored proteostasis (Fig. 1F and fig. S2, F to H). Similar to *iKras*^{R} cells, continuous sotorasib treatment was required to maintain the resistance of MIA-PaCa-2^{R} cells to KRASi (fig. S2, I and J). MIA-PaCa-2 xenograft tumors were initially sensitive to sotorasib treatment but relapsed after 6 weeks (fig. S2K). PROTEOSTAT staining revealed increased protein aggregation after initial sotorasib treatment and restored proteostasis in relapsed tumors (Fig. 1, G and H). These data demonstrate that oncogenic KRAS inhibition induces protein aggregation and severely disrupts proteostasis. However, KRASi-resistant cancer cells able to grow in the absence of mutant KRAS gain the capacity to overcome associated proteotoxic stress and regain proteostasis.

Oncogenic KRAS inactivation differentially affects the key nodes of proteostasis regulatory network

Proteostasis is maintained by an integrated network that includes translation, protein quality control mechanisms that regulate the content and quality of the proteome, and protein degradation pathways, such as the ubiquitin-proteasome system and the autophagy-lysosome system, which eliminate misfolded or aggregated proteins (*42*, *43*, *58*). *Kras*^{G12D} inactivation through Dox withdrawal in *iKras* cells did not alter overall proteasomal activity (fig. S3, A and B). We observed increased LC3 cleavage and lysosome-associated membrane protein type 1 (LAMP1) levels upon *Kras*^{G12D} inactivation (fig. S3C), indicating elevated autophagy and lysosomal activities, which unlikely caused increased protein aggregation. Next, we examined the protein quality control and stress response pathways that monitor and regulate protein folding, including the UPR and HSR. *Kras*^{G12D} inactivation in *iKras*^{P} cells markedly reduced IRE1α protein levels, *Xbp1* splicing, and the expression levels of the XBP1s targets *Edem1* and *Sec61a1* (Fig. 2A and fig. S3, D to F). As a control, Dox treatment had no impacts on IRE1α levels in the constitutively activated *LSL-Kras*^{G12D} cells (fig. S3G). By contrast, ATF6 was barely affected, and phospho-PERK was negatively correlated with IRE1α/XBP1 in response to *Kras*^{G12D} inactivation (Fig. 2A). Resolving protein aggregation with tauroursodeoxycholic acid (TUDCA), a chemical chaperone that promotes protein folding and stimulates molecular chaperone function (*59*, *60*), completely blunted *Kras*^{G12D} inactivation-induced phospho-PERK (fig. S3, H to J), indicating that PERK was activated as a result of the disrupted proteostasis by KRASi. PERK is one of the four kinases [general control nonderepressible 2 (GCN2), PERK, heme-regulated inhibitor (HRI), and RNA-activated protein kinase (PKR)] that phosphorylate eIF2α and regulate the integrated stress response (ISR) (*61*). Unexpectedly, eIF2α and its downstream ATF4 followed an opposite pattern as that of phospho-PERK upon *Kras*^{G12D} inactivation (Fig. 2A). *Perk* deletion had no effects on their levels (fig. S3K). Genetic deletion of each ISR kinases demonstrated that phospho-eIF2α and ATF4 were dependent on GCN2, which was activated in *iKras*^{P} cells (Fig. 2A and fig. S3K). These data establish GCN2 as a regulator of ISR in this context. IRE1α/XBP1, but not GCN2 or eIF2α, was selectively restored in *iKras*^{R} cells (Fig. 2A and fig. S3, D to F). Consistently, sotorasib treatment reduced IRE1α protein levels in MIA-PaCa-2 and H358 cells, but IRE1α was restored after the acquisition of sotorasib resistance (Fig. 2, B and C, and fig. S2, D and E). IRE1α immunostaining in the parental, *Kras*^{G12D}-extincted, relapsed *iKras* GEMM tumors (Fig. 2D) and sotorasib-treated MIA-PaCa-2 human xenograft tumors (Fig. 2E) all independently confirmed the in vivo pattern of acute IRE1α suppression followed by reactivation in relapsed tumors. We also confirmed the continuous in vivo suppression of phospho-GCN2, phospho-eIF2α, and ATF4 through sotorasib treatment in MIA-PaCa-2 xenograft tumors (fig. S3L). Collectively, these data show that acute oncogenic KRAS inactivation inhibits the IRE1α branch of the UPR and GCN2-regulated ISR. However, only IRE1α is reactivated in the KRASi-resistant tumors.

In contrast with our findings on IRE1α, sotorasib continuously suppressed HSF1 phosphorylation at serine residue 326 (S326), which is required for HSF1 activation (*62*), in both parental and sotorasib-resistant MIA-PaCa-2 and H358 cells (Fig. 2, B and C). The inhibitory phosphorylation of HSF1 at S121 (*62*) was increased in both sotorasib-resistant MIA-PaCa-2^{R} and H358^{R} cells (Fig. 2, B and C). As a result, sotorasib treatment markedly reduced the expression of HSF1 target genes, including *HSPA6* and *HSPA1B*, in both parental and sotorasib-resistant MIA-PaCa-2 and H358 cells (fig. S3, M and N). We confirmed the in vivo suppression of HSF1-S326 phosphorylation by means of sotorasib treatment in MIA-PaCa-2 human xenograft tumors (Fig. 2E) and marked reduction of HSF1 luciferase reporter activities in sotorasib-resistant MIA-PaCa-2^{R} cells (fig. S3, O and P). The genetic inactivation of *Kras*^{G12D} also resulted in the considerable reduction of HSF1 luciferase reporter activities in both *iKras*^{P} and *iKras*^{R} cells (Fig. 2F). Collectively, these data demonstrate that oncogenic KRAS inactivation initially impairs both the ER and cytosolic protein quality control machinery as well as the ISR. However, only IRE1α/XBP1, but not HSF1 or ISR, is restored in the KRASi-resistant tumors.

IRE1α/XBP1 is required for maintaining proteostasis in KRASi-resistant cancer cells

Next, we examined the necessity of IRE1α/XBP1 for proteostasis maintenance. In *iKras*^{P} cells, *Ire1α* and *Xbp1* depletion modestly induced protein aggregation (Fig. 2, G and H, and fig. S4A), with minimal effects on cell growth (Fig. 2I). By contrast, *Ire1α* and *Xbp1* depletion in the *iKras*^{R} cells led to profound protein aggregation and PERK phosphorylation (Fig. 2, G and H, and fig. S4, A to F) and significantly inhibited cell growth (Fig. 2I). These effects were rescued by TUDCA (Fig. 2, G to I, and fig. S4B), demonstrating the importance of IRE1α/XBP1 in the maintenance of proteostasis and cell survival. Depletion of *Perk* did not affect proteostasis and cell growth of *iKras*^{R} cells and had no impacts on *Ire1α* depletion–induced phenotypes (fig. S4, G to K), suggesting that PERK is dispensable for KRASi-resistant cancer cells. The kinase activity of IRE1α is required for its RNase activation (*63*). In contrast to wild-type (WT) IRE1α, neither kinase-dead IRE1α^{K599A} nor RNase-dead IRE1-α^{K907A} mutant (*63*) was able to rescue *Ire1α* depletion–induced phenotypes in *iKras*^{R} cells (fig. S5, A to D), suggesting that IRE1α's function depends on its catalytic RNase activity. In addition to XBP1s, IRE1α RNase also cleaves ER-localized RNAs through the regulated IRE1α-dependent decay (RIDD) pathway (*64*, *65*). Although some RIDD targets were regulated by IRE1α (fig. S5, E and F), restoration of XBP1s completely rescued the *Ire1α* depletion–induced protein aggregation and cell growth defects in *iKras*^{R} cells (fig. S5, G to I). These data establish that IRE1α RNase activity and RNase-dependent *XBP1* splicing drives proteostasis in KRASi-resistant cancer cells. Consistently, CRISPR-mediated *Ire1α* or *Xbp1* deletion (fig. S5J) resulted in more severe protein aggregation in *iKras*^{R} cells than in *iKras*^{P} cells (fig. S5K). Taken together, we demonstrate that IRE1α/XBP1 is indispensable for the maintenance of balanced proteostasis in KRASi-resistant cancer cells.

The MAPK pathway regulates IRE1α/XBP1 in parental KRAS-driven cancer cells

We aimed to determine the mechanism through which oncogenic KRAS regulates IRE1α. MAPK and PI3K are two major effector pathways downstream of oncogenic KRAS (*6*). The MEK inhibitor trametinib and the ERK inhibitor SCH772984 both substantially reduced IRE1α protein levels in *iKras*^{P} cells (Fig. 3A). By contrast, the PI3K inhibitor pictilisib and the AKT

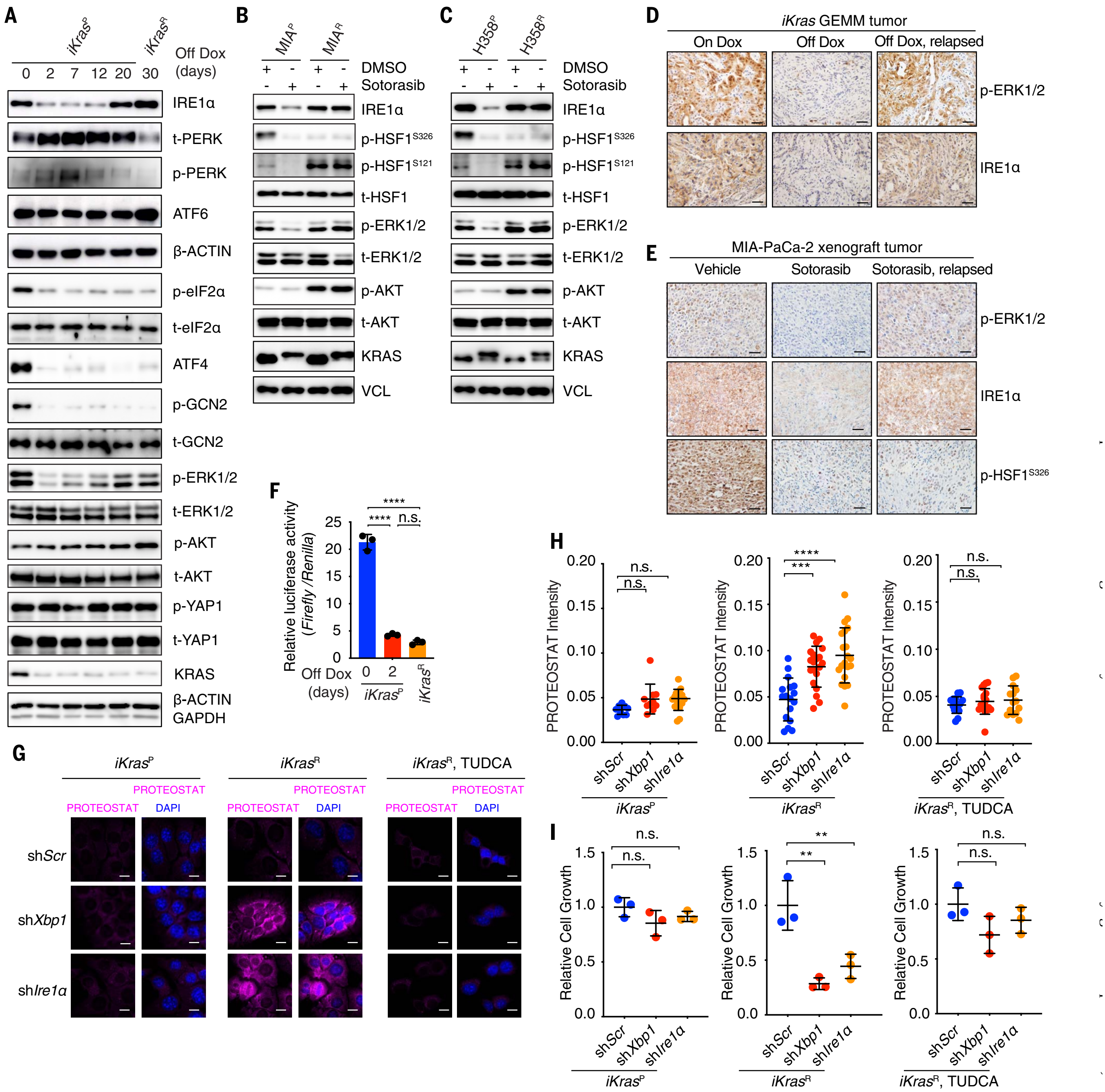


Fig. 2. Oncogenic KRAS inactivation differentially affects the key nodes of the proteostasis regulatory network. (**A**) Immunoblot with indicated antibodies in whole-cell lysates of *iKras*^{P} at different time points after *Kras*^{G12D} inactivation through Dox-withdrawal (Off Dox) until the cells acquired resistance to *Kras*^{G12D} inactivation (*iKras*^{R} cell). (**B** and **C**) Immunoblot with indicated antibodies in whole-cell lysates of parental or sotorasib-resistant (B) MIA-PaCa-2 or (C) H358 cells treated with DMSO or 30 nM sotorasib. (**D**) Immunohistochemical (IHC) staining with indicated antibodies in *iKras* GEMM tumors treated with doxycycline (On Dox), Dox withdrawal for 3 days (Off Dox), or relapsed after 30 weeks of Dox-withdrawal (Off Dox, relapsed). (**E**) IHC staining with indicated antibodies in MIA-PaCa-2 xenograft tumors treated with vehicle, sotorasib (30mg/kg for 1 day), or relapsed after 9 weeks of sotorasib treatment (30 mg/kg). (**F**) Relative HSE luciferase activity in *iKras*^{P} or *iKras*^{R} cells cultured in the presence or absence of Dox for 2 days. Data are shown as mean ± SD, n = 3 biological replicates. (**G**) Representative images and (**H**) quantification of PROTEOSTAT (magenta) and DAPI (blue) staining in *iKras*^{P} or *iKras*^{R} cells infected with lentiviruses encoding *scramble* short hairpin RNA (shRNA) (sh*Scr*), *Xbp1* shRNA (sh*Xbp1*), or *Ire1α* shRNA (sh*Ire1α*). Cells were treated with 2.5 mM TUDCA dissolved in water for 2 days as indicated. Data represent the average fluorescence intensity of PROTEOSTAT per cell from each image acquired and presented as mean ± SD from n = 10 (On Dox), n = 17 (Off Dox), or n = 17 (Off Dox + TUDCA) images. (**I**) CCK-8 assay was used to quantify cell viability of *iKras* cells treated as in (G) and (H). Data are presented as mean ± SD relative to sh*Scr*, n = 3 biological replicates. Ordinary one-way ANOVA with Dunnett's multiple comparisons test in (H) and (I) and ordinary one-way ANOVA with Tukey's multiple comparisons test in (F) was used to calculate P values. n.s., not significant, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. Scale bar, (D) and (E), 40 μm; (G) 20 μm.

inhibitor MK2206 had little impact on IRE1α levels in *iKras*[P] cells (Fig. 3B). Examination of five additional *KRAS*-mutant cell lines confirmed that MEK/ERK inhibitors, but not PI3K/AKT inhibitors, reduced IRE1α protein levels (fig. S6, A to C). These effects were similar to what was observed in response to the genetic or pharmacological inhibition of oncogenic KRAS (Fig. 3F and fig. S6D). Recent studies show that SHP2 (SH2 containing protein tyrosine phosphatase-2) is critical for $KRAS^{G12C}$ cycling and ERK activation (*30*, *33*, *66*–*69*). Inhibition of SHP2 with SHP099 substantially suppressed ERK activity and reduced IRE1α levels in *KRAS*^{G12C}-mutant H358 cells (fig. S6E), whereas modest effect was observed in *KRAS*^{G12D}-mutant *iKras* cells because of the limited intrinsic guanosine triphosphatase (GTPase) activity of $KRAS^{G12D}$ (fig. S6F). SHP099 also inhibited the growth of sotorasib-resistant $H358^R$ cells and MIA-PaCa-2^R cells (fig. S6, G and H). Long-term treatment of H358 cells with SHP099 led to drug resistance and recovered proteostasis (fig. S6, I to K), accompanied with IRE1α restoration in the resistant cells (fig. S6L). Depletion of *IRE1α* in SHP2i-resistant cells resulted in marked protein aggregation and resensitized the cells to SHP099 (fig. S6, M to O). Another upstream activator of RAS signaling is epidermal growth factor receptor (EGFR); suppression of EGFR with gefitinib markedly reduced ERK and IRE1α levels in H358 cells (fig. S6P). By contrast, inhibition of MEK/ERK in nonmalignant BEAS-2B lung epithelial cells with nononcogenic RAS signaling barely affected IRE1α levels (fig. S6, Q and R). Using tissue microarrays, we found that IRE1α levels correlated with phospho-ERK levels in treatment-naïve PDAC patient samples (Fig. 3, C and D), and high expression of IRE1α was associated with higher histologic tumor grade (Fig. 3E). Collectively, these data demonstrate that oncogenic KRAS-mediated MAPK pathway activation leads to the activation of IRE1α in parental *KRAS*-mutant cancers.

MAPK promotes IRE1α protein stability by inhibiting SEL1L/HRD1-mediated IRE1α ubiquitination

Sotorasib treatment did not down-regulate *IRE1α* mRNA levels (fig. S7A), but it considerably reduced IRE1α protein abundance in H358 cells (Fig. 3F). Treatment with the proteasome inhibitor MG132 rescued both sotorasib- and trametinib-induced reductions in IRE1α protein levels in H358 cells (Fig. 3F). Similarly, the observed reduction in IRE1α protein levels in response to *Kras*^{G12D} inactivation, trametinib, or SCH772984 treatment could be rescued by MG132 in *iKras*[P] cells (fig. S7, B and C). These data demonstrate that the inhibition of oncogenic KRAS or MAPK promotes proteasome-mediated IRE1α degradation in parental *KRAS*-mutant cancers.

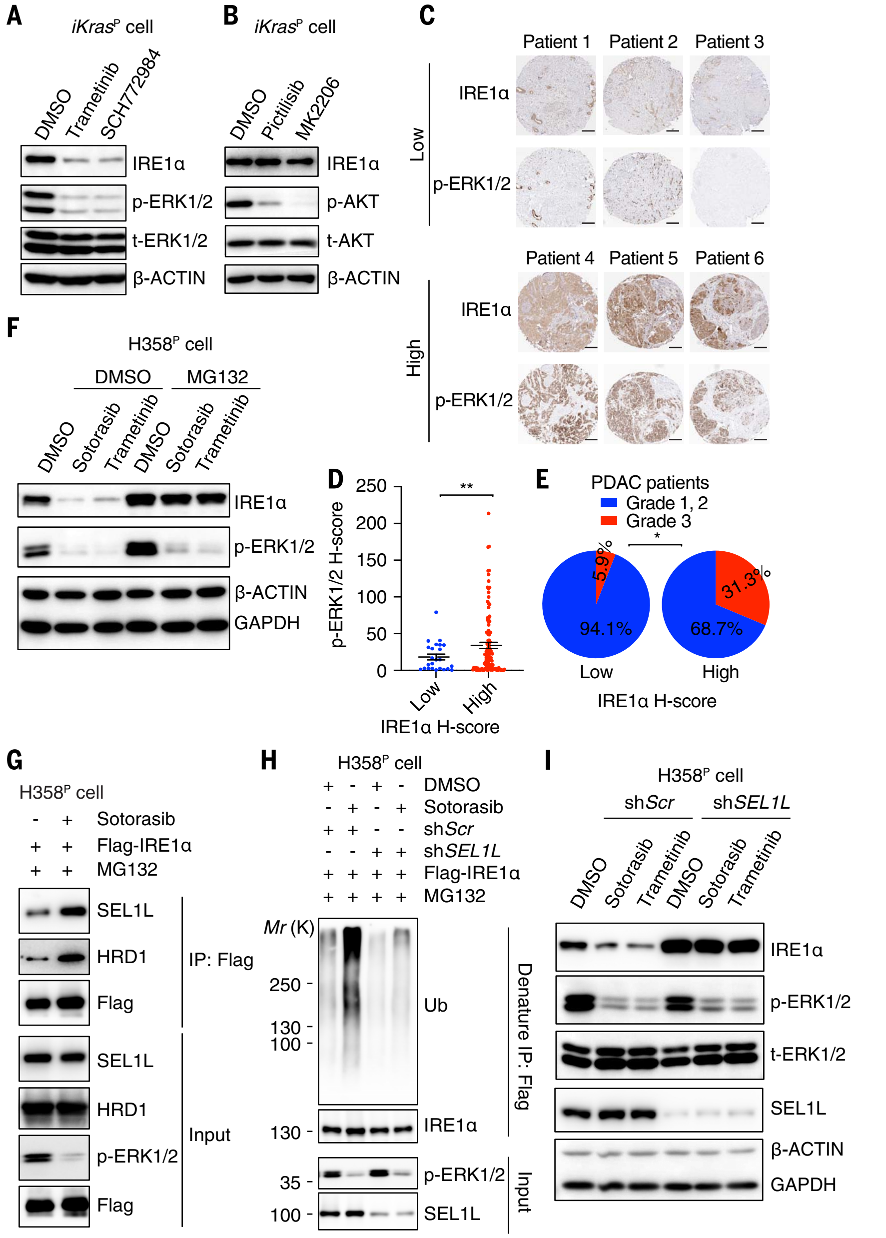


Fig. 3. KRAS-MAPK signaling stabilizes IRE1α through inhibiting SEL1L-HRD1 mediated IRE1α ubiquitination. (**A** and **B**) Immunoblot with indicated antibodies in whole-cell lysates of *iKras*[P] cells treated with DMSO, trametinib (MEK inhibitor; 20 nM), SCH772984 (ERK inhibitor; 1 μM), pictilisib (PI3K inhibitor; 1 μM), or MK2206 (AKT inhibitor; 2 μM) as indicated for 2 days. (**C**) Representative images of IHC staining of IRE1α and p-ERK1/2 in tissue microarray of treatment naïve tumors from PDAC patients. Scale bar, 200 μm. (**D**) H-score of p-ERK1/2 in tissue microarray samples with distinct IRE1α intensities. Data are presented as mean ± SEM. (**E**) Proportion of patients with different tumor grades in PDAC patients with low or high IRE1α H-score. (**F**) Immunoblot with indicated antibodies in whole-cell lysates of $H358^P$ cells treated with DMSO, 30 nM sotorasib, or 20 nM trametinib for 2 days. Cells were treated with DMSO or 1 μM MG132 for 12 hours before harvest. (**G**) Sotorasib promotes the interaction between IRE1α and SEL1L/HRD1. $H358^P$ cells expressing Flag-IRE1α were treated with DMSO or 30 nM sotorasib for 2 days and subjected to immunoprecipitation (IP) with anti-Flag M2 agarose beads. (**H**) Sotorasib promotes SEL1L-dependent IRE1α ubiquitination. $H358^P$ cells expressing Flag-IRE1α and sh*Scr* or sh*SEL1L* were treated with DMSO or 30 nM sotorasib for 2 days and subjected to denature IP with anti-Flag M2 agarose beads. The immunoblot was probed with antibody to ubiquitin (Ub) to detect IRE1α ubiquitination. In (G) and (H), MG132 (1 μM) was added into the culture medium 12 hours before harvest. (**I**) Immunoblot with indicated antibodies in whole-cell lysates of $H358^P$ cells infected with lentiviruses encoding sh*Scr* or sh*SEL1L* and treated with DMSO, 30 nM sotorasib, or 20 nM trametinib for 2 days. Two-tailed, unpaired Student's *t* test with Welch's correction in (D) and Fisher's exact test in (E) was used to calculate *P* values. *$P < 0.05$, **$P < 0.01$.

IRE1α is a bona fide substrate of the SEL1L/HRD1 ER-associated degradation (ERAD) complex (*70*), which is composed of the E3 ubiquitin ligase HRD1 and its adapter protein SEL1L (suppressor/enhancer of lin-12-like) (*71*). The SEL1L/HRD1 complex ubiquitinates and targets IRE1α for proteasomal degradation in multiple cell types (*70*, *72*, *73*). Sotorasib treatment substantially enhanced the association between IRE1α and the SEL1L/HRD1 complex (Fig. 3G), resulting in increased IRE1α ubiquitination in H358 cells (Fig. 3H). Sotorasib also promoted the interaction of IRE1α with p97 and NPL4 (nuclear protein localization protein 4) (fig. S7, D and E), which deliver the ubiquitinated ERAD substrates to the proteasome for degradation (*74*). *SEL1L* depletion reduced sotorasib- or trametinib-induced IRE1α ubiquitination and restored IRE1α protein levels (Fig. 3, H and I, and fig. S7F), leading to the prevention of KRAS-MAPK inhibition–induced protein aggregation (fig. S7G). Inhibition of p97 with CB5083 (*75*) also rescued sotorasib-induced IRE1α degradation (fig. S7I). Consistently, *Sel1l* or *Hrd1* depletion in *iKras*^{P} cells blocked the induction of IRE1α degradation and prevented protein aggregation in response to *Kras*^{G12D} extinction and trametinib or SCH772984 treatment (fig. S7, J to P). These data demonstrate that oncogenic KRAS-mediated MAPK activation stabilizes IRE1α protein by preventing SEL1/HRD1-mediated ubiquitination and subsequent proteasomal degradation.

ERK directly interacts with and phosphorylates IRE1α

Oncogenic KRAS-MAPK did not directly regulate ERAD complex expression in *iKras*^{P} or H358 cells (Fig. 3I and fig. S7, J, M, and Q). It is well established that phosphorylation often interferes with protein-protein interactions and thus regulates protein ubiquitination and stability. We tested whether MAPK might promote IRE1α phosphorylation, resulting in IRE1α dissociation from the SEL1L/HRD1 complex. The expression of a constitutively activated MEK construct (MEK^{DD}) dramatically enhanced IRE1α phosphorylation in 293T cells detected with an antibody to phospho-MAPK substrate motif (fig. S8A). By contrast, sotorasib treatment substantially reduced IRE1α phosphorylation levels compared with control H358 cells (Fig. 4A). Co-immunoprecipitation (co-IP) assay demonstrated that ERK interacted with IRE1α in 293T cells and H358 cells (Fig. 4B and fig. S8B). Furthermore, glutathione *S*-transferase (GST) pull-down and in vitro kinase assays confirmed that both ERK1 and ERK2 directly interacted with and phosphorylated IRE1α in vitro (Fig. 4, C and D, and fig. S8, C to H). Depletion of *ERK1* or *ERK2* in H358 cells revealed that both ERK1 and ERK2 regulated IRE1α phosphorylation (fig. S8I). IRE1α possesses three putative ERK binding D-motifs (Fig. 4E) (*76*). Deletion of the D-motif at amino acids 687 to 701 largely disrupted the binding between ERK and IRE1α (fig. S8J). Collectively, these data demonstrate that ERK directly interacts with and phosphorylates IRE1α.

Sequence analysis showed that human IRE1α contains four putative ERK phosphorylation sites at Ser^{525} (S525), Ser^{529} (S529), Ser^{549} (S549), and Thr^{973} (T973), which is consistent with the minimal ERK substrate motif pS/T*P (Fig. 4E). Mass spectrometry analysis of IRE1α protein purified from control or MEK^{DD}-expressing 293T cells confirmed the ERK-dependent phosphorylation of IRE1α at S525, S529, S549, and T973 (fig. S9, A to C). We mutated the identified phospho-serine or -threonine amino acids to alanine (A) and performed an in vitro kinase assay using [γ-^{32}P] adenosine 5′-triphosphate (ATP). The kinase-dead form of IRE1α^{K599A} was used as a backbone to exclude the effects of IRE1α autophosphorylation. ERK was able to directly phosphorylate kinase-dead autophosphorylation-deficient IRE1α^{K599A} (Fig. 4F). The T973A mutation (in which threonine was replaced with alanine at position 973) markedly reduced but did not eliminate ERK-dependent IRE1α^{K599A} phosphorylation (Fig. 4F). Additional mutation of S525, S529, or S549 to A together with T973A further decreased the IRE1α phosphorylation (Fig. 4F). The simultaneous mutation of all four sites largely abolished the ERK-dependent IRE1α^{K599A} phosphorylation (Fig. 4F). In agreement, mutation of these four S or T to A (designated as 4A mutant) diminished ERK-dependent phosphorylation on WT IRE1α (Fig. 4G and fig. S9D). The IRE1α mutations did not affect IRE1α binding with ERK (fig. S9, D and E). Collectively, these data identify S525, S529, S549, and T973 as ERK phosphorylation sites on IRE1α. Analysis of Clinical Proteomic Tumor Analysis Consortium (CPTAC) datasets in patients with NSCLC (*77*) showed a statistically significant correlation between IRE1α-S549 phosphorylation and ERK phosphorylation in treatment-naïve NSCLC patients (Fig. 4H and fig. S9F). The peptides containing S525, S529, and T973 were not covered in the CPTAC datasets and could not be evaluated in patients.

To determine the functional relevance of IRE1α phosphorylation sites, we generated a loss-of-function mutation for each site. Single-site mutation was insufficient to promote IRE1α interaction with HRD1 and did not alter IRE1α levels in *iKras* cells (Fig. 4, I and J). Simultaneous mutation of all four sites promoted IRE1α interaction with HRD1 and its degradation (Fig. 4, I and J). In vitro pull-down assays confirmed that ERK-mediated IRE1α phosphorylation reduced IRE1α interaction with HRD1/SEL1L and that phospho-deficient IRE1α^{4A} mutant bound to HRD1 regardless of ERK presence (fig. S10, A to E). By contrast, gain-of-function phosphomimetic mutation for each individual site (S525D, S529D, S549D, or T973E, where D is aspartate and E is glutamate) disrupted IRE1α interaction with HRD1 (Fig. 4K), leading to the stabilization of IRE1α protein in the absence of ERK (Fig. 4L). Similar effects were observed for IRE1α^{SDTE} mutant with gain-of-function mutation for all four sites (Fig. 4, K and L, and fig. S10F). IRE1α^{SDTE} was resistant to sotorasib- or SCH772984-promoted protein degradation in MIA-PaCa-2 cells (fig. S10, G and H). As a result, sotorasib failed to induce protein aggregation in IRE1α^{SDTE}-expressing MIA-PaCa-2 tumors (Fig. 5, A to C). These tumors became partially resistant to sotorasib-induced antitumor effects (Fig. 5D). In line with these data, the IRE1α^{SDTE} mutant rescued *IRE1α* depletion–induced protein aggregation, phospho-PERK, and cell growth defects in *iKras*^{R} cells (fig. S10, I to L). Single-site phospho-deficient IRE1α mutant also rescued these phenotypes owing to the presence of the other three phosphorylated sites (fig. S10, L to O). The phospho-deficient IRE1α^{4A} mutant failed to restore IRE1α protein levels and was unable to rescue these phenotypes (fig. S10, L to O). Collectively, these data demonstrate that IRE1α phosphorylation at S525, S529, S549, and T973 inhibits IRE1α association with the ERAD complex, leading to enhanced stability, maintaining proteostasis.

A screen of 32 serine and threonine phosphatases in 293T cells revealed that expression of SCP3 substantially reduced MEK^{DD}-induced IRE1α phosphorylation (fig. S11, A to C). In vitro phosphatase assays and co-IP experiments confirmed that SCP3 interacted with and directly dephosphorylated IRE1α (Fig. 5E, and fig. S11D). *Scp3* silencing increased IRE1α phosphorylation, and overexpression of SCP3 reduced IRE1α phosphorylation, in *iKras* cells (Fig. 5, F and G). Similarly, *SCP3* deletion slowed down sotorasib-induced IRE1α dephosphorylation in H358 cells (fig. S11E). These data identified SCP3 as the phosphatase regulating IRE1α phosphorylation, although KRAS did not directly alter SCP3 levels or activities (fig. S11, F to I). Collectively, these analyses establish a mechanism of IRE1α regulation by oncogenic KRAS.

Multiple pathways converge to reactivate IRE1α in KRASi-resistant cancer cells

Next, we sought to determine how IRE1α evades oncogenic KRAS inhibition and determine the reactivation mechanism in KRASi-resistant cells. Oncogenic KRAS was efficiently suppressed by Dox withdrawal in *iKras*^{R} cells (Fig. 2A and fig. S1B), and most KRAS proteins were bound by sotorasib in sotorasib-resistant H358 (H358^R) and MIA-PaCa-2 (MIA-PaCa-2^R) cells, similar to observations of parental cells (Fig. 2, B and C). Furthermore, silencing of *KRAS* in H358^R cells did not hinder IRE1α reactivation (fig. S12A). These data exclude the possibility that the inefficient inhibition of oncogenic KRAS drives IRE1α reactivation in these

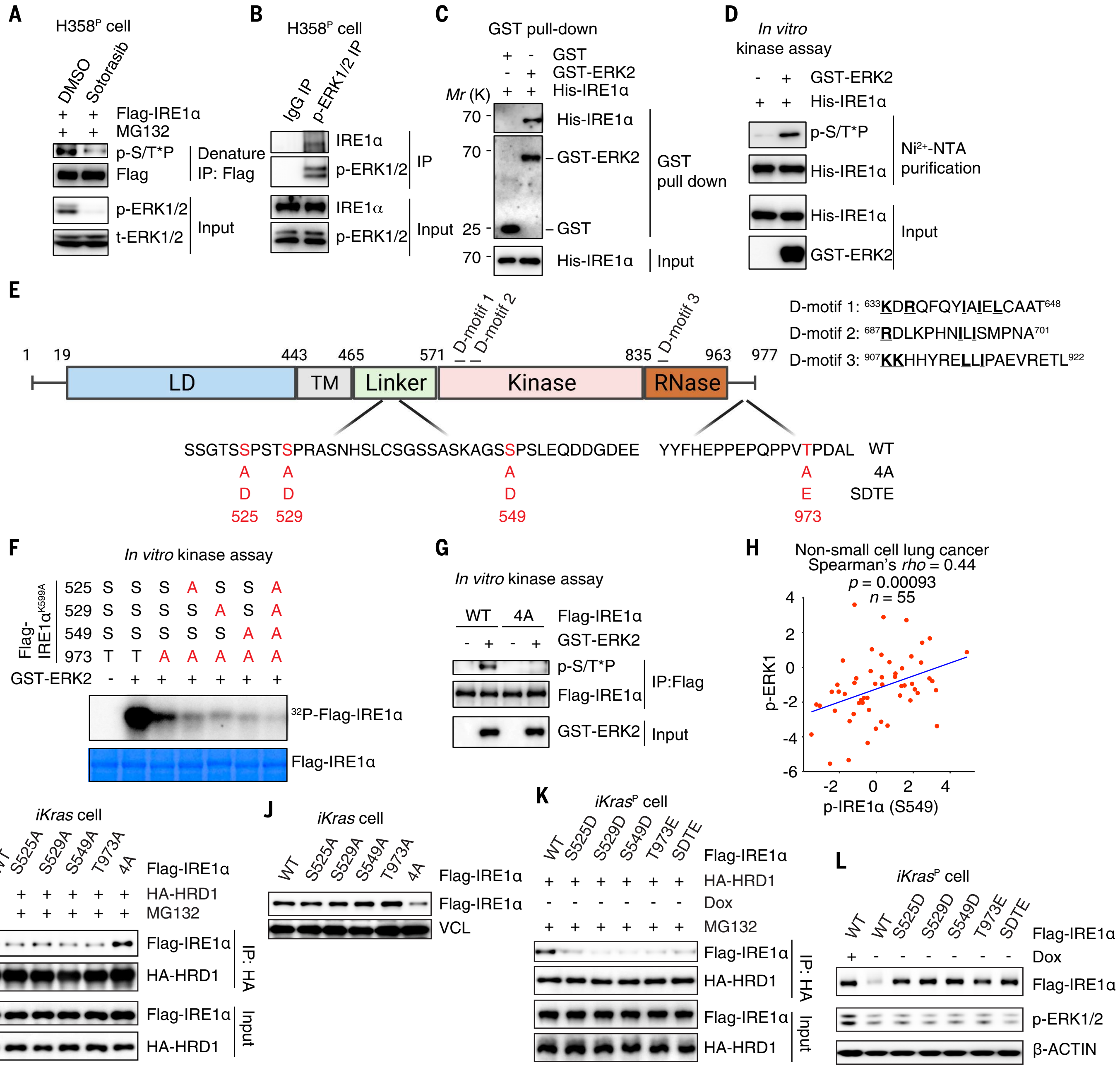


Fig. 4. ERK directly phosphorylates and stabilizes IRE1α. (**A**) H358[P] cells expressing Flag-IRE1α were treated with DMSO or 30 nM sotorasib for 2 days and subjected to denature IP with anti-Flag M2 agarose beads. The immunoblot was probed with anti-phospho-MAPK substrates motif (S/T*P) antibody (p-S/T*P) to detect IRE1α phosphorylation. (**B**) Whole-cell lysate of H358[P] cells were subjected to co-IP with rabbit anti-p-ERK1/2 antibody or normal rabbit immunoglobulin G (IgG). (**C**) GST pull-down assay was performed by using recombinant His-tagged IRE1α and GST-ERK2 protein. (**D**) In vitro kinase assay was performed by using recombinant GST-ERK2 and His-IRE1α. After phosphorylation reaction, the proteins were denatured with 8M urea buffer and subjected to purification of His-IRE1α with Ni^{2+}-NTA agarose to detect IRE1α phosphorylation by use of anti-p-S/T*P antibody. (**E**) Schematic illustration of human IRE1α protein domains, three putative ERK binding D-motifs, and ERK phosphorylation sites at Ser^{525}, Ser^{529}, Ser^{549}, and Thr^{973} (red). Phospho-deficient (4A) and phospho-mimetic (SDTE) mutations of IRE1α are shown. LD, luminal domain; TM, transmembrane domain. (**F**) In vitro [γ-^{32}P] ATP kinase assay by using different Flag-tagged IRE1α mutants and GST-ERK2. The IRE1α phosphorylation was detected with autoradiography. One-Step Blue Protein Stain (Biotium) was used to detect IRE1α protein loading. (**G**) In vitro kinase assay by using equal amount of Flag-tagged WT or phospho-deficient IRE1α mutant proteins (4A) and GST-ERK2. (**H**) Spearman correlation between p-IRE1α (at S549) and p-ERK1 (at Y204) in 55 patients with NSCLC. (**I** and **K**) Whole-cell lysates of *iKras* cells expressing HA-HRD1 together with WT or mutant IRE1α cultured in the absence of Dox were subjected to IP with anti-HA agarose beads to detect IRE1α interaction with HRD1. In (A), (I), and (K), MG132 (1 μM) was added into the culture medium 12 hours before harvest. (**J** and **L**) Immunoblot of WT or mutant IRE1α in whole-cell lysates of *iKras* cells cultured in the presence or absence of Dox for 2 days.

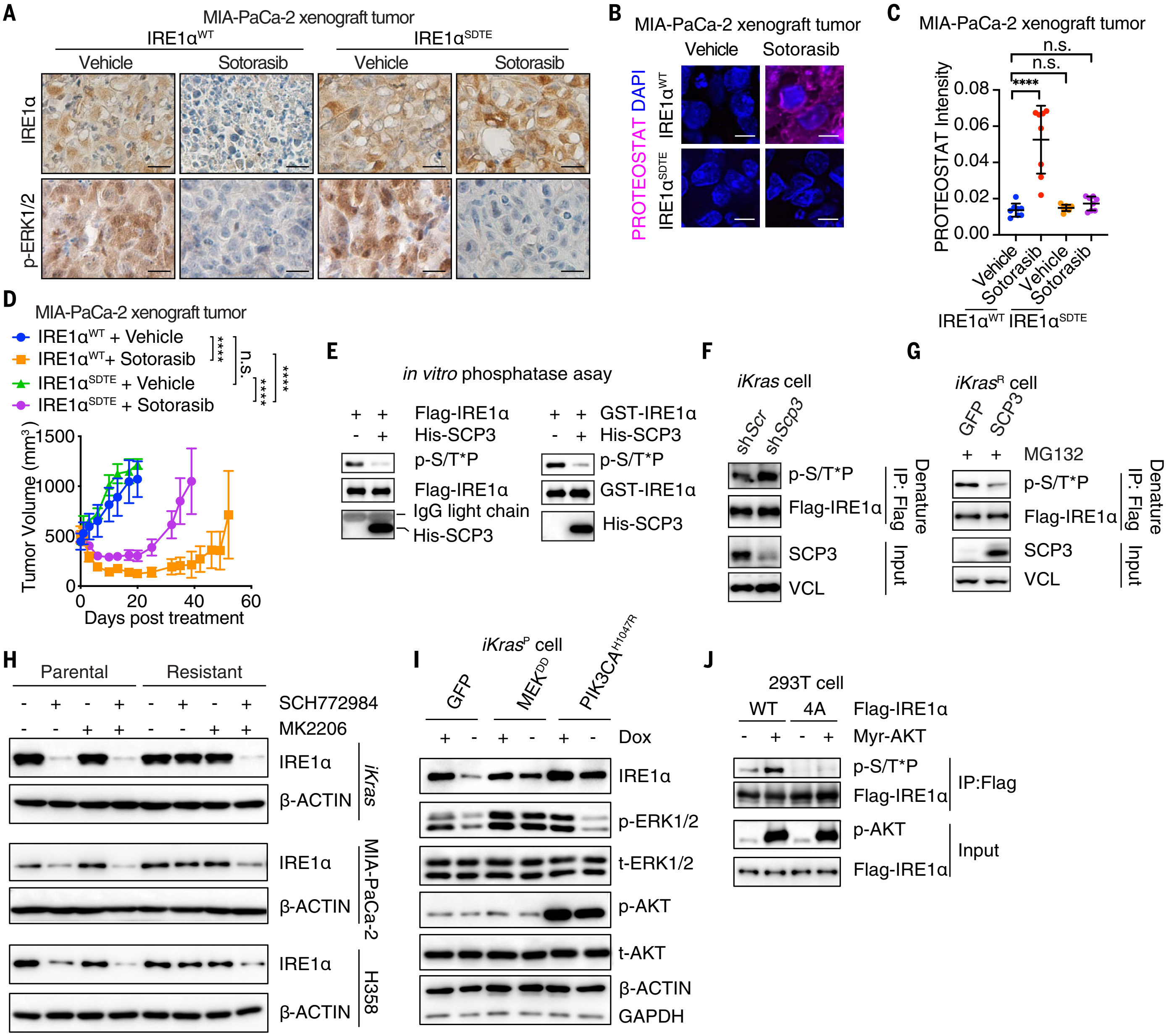


Fig. 5. Multiple pathways converge to restore IRE1α in KRASi-resistant cancer cells. (**A**) IHC staining of p-ERK1/2 and IRE1α in shRNA-resistant *IRE1α*^{WT}- or *IRE1α*^{SDTE}-transduced, endogenous *IRE1α*-depleted MIA-PaCa-2 tumors treated with vehicle or sotorasib (100 mg/kg) for 4 days. Scale bar, 40 μm. (**B**) Representative images and (**C**) quantification of PROTEOSTAT (magenta) and DAPI (blue) staining in MIA-PaCa-2 tumors as in (A). Data represent average fluorescence intensity of PROTEOSTAT per cell from each image acquired and are presented as mean ± SD from n = 8 independent images. Scale bar, 20 μm. (**D**) Tumor volume quantification of MIA-PaCa-2 tumors as in (A). (**E**) In vitro phosphatase assay. (Left) Phosphorylated Flag-IRE1α protein purified from MEK^{DD}-expressing 293T cells or (right) recombinant IRE1α protein phosphorylated by recombinant ERK2 in vitro were subjected to in vitro phosphatase assay with recombinant SCP3. (**F** and **G**) *iKras* cells expressing Flag-IRE1α were infected with shRNA targeting (F) *Scr* or *Scp3* and (G) GFP- or SCP3- expressing lentivirus. The whole-cell lysates were subjected to denature IP with anti-Flag M2 agarose beads, followed by immunoblot with antibody to pS/T*P to detect IRE1α phosphorylation. (**H**) Immunoblot of IRE1α in whole-cell lysates of parental and KRAS inhibition-resistant cells treated with DMSO, 2 μM MK2206, and/or 1 μM SCH772984 for 2 days (MIA-PaCa-2 and H358 cells) or 14 days (*iKras* cells). (**I**) Immunoblot of IRE1α in whole-cell lysates of *iKras*^{P} cells expressing GFP, MEK^{DD}, or PIK3CA^{H1047R} in the presence or absence of Dox for 2 days. (**J**) 293T cells expressing IRE1α^{WT} or IRE1α^{4A} in the presence or absence of myr-AKT were subjected to IP with anti-Flag M2 agarose beads followed by immunoblot to detect IRE1α phosphorylation. Ordinary one-way ANOVA with Dunnett's multiple comparisons test in (C) and two-way ANOVA with Bonferroni's multiple comparisons test in (D) were used to calculate P values. n.s., not significant; $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

resistant cells. eIF2α phosphorylation inhibits global protein synthesis (*61*, *78*). Consistent with the inactivated phospho-eIF2α in KRASi-resistant cells (Fig. 2A), we observed increased global protein synthesis in *iKras*^{R} cells evidenced by enhanced puromycin incorporation compared with that in *iKras*^{P} cells (fig. S12B). However, inhibition of protein synthesis with cyclohexamide did not affect IRE1α/XBP1 in *iKras*^{R} cells (fig. S12, C and D). Furthermore, ER stress sensing–deficient IRE1α$^{\Delta 2M}$ mutant (*79*) was similarly restored in KRASi-resistant cells as that of WT IRE1α (fig. S12, E and F) and successfully rescued *IRE1α* depletion–induced phenotypes (fig. S12, G to J). These data

demonstrate that IRE1α is reactivated in KRASi-resistant cells in an ER stress–independent manner.

Recent studies report reactivated ERK and AKT as sotorasib-resistant mechanisms in patients (*18*, *19*, *21*–*23*). We observed the reactivation of phospho-ERK and the hyperactivation of phospho-AKT in sotorasib-resistant MIA-PaCa-2^R and $H358^R$ cells compared with their respective parental cells (Fig. 2, B and C). Unexpectedly, the inhibition of reactivated ERK through SCH772984 treatment was insufficient to suppress IRE1α levels in MIA-PaCa-2^R, $H358^R$, and *iKras*^{R} cells (Fig. 5H) as well as in MIA-PaCa-2^R and *iKras*^{R} tumors in vivo (fig. S13, A and B). Similarly, the suppression of AKT through MK2206 treatment had no effect on IRE1α levels in KRASi-resistant cells (Fig. 5H). However, the simultaneous suppression of both ERK and AKT successfully blunted IRE1α reactivation in MIA-PaCa-2^R, $H358^R$, and *iKras*^{R} cells (Fig. 5H) and in MIA-PaCa-2^R and *iKras*^{R} tumors in vivo (fig. S13, A and B). Consistently, the hyperactivation of either the MEK/ERK pathway, through expression of MEK^{DD}, or the PI3K/AKT pathway, through the expression of constitutively active $PIK3CA^{H1047R}$ or myr-AKT, resulted in IRE1α restoration in the absence of oncogenic KRAS in *iKras* cells (Fig. 5I and fig. S13C). myr-AKT also promoted WT, but not phospho-deficient 4A mutant, IRE1α phosphorylation at serine and threonine residues (Fig. 5J), suggesting that these phosphorylation sites are regulated by both ERK and hyperactivated AKT in KRASi-resistant cells. In agreement, either the activation of MEK/ERK, through MEK^{DD} expression, or the hyperactivation of AKT, through myr-AKT expression, was sufficient to disrupt the interaction of the SEL1L/HRD1 E3 ligase complex with IRE1α (fig. S13D). By contrast, the simultaneous suppression of both ERK and AKT, but not ERK or AKT alone, promoted IRE1α interaction with the SEL1L/HRD1 E3 ligase complex in sotorasib-resistant MIA-PaCa-2^R cells (fig. S13E). Yes-associated protein 1 (YAP1) also drives resistance of certain tumors to KRASi (*37*, *38*). However, deletion of *Yap1* in *iKras*^{R_YAP1} cells derived from YAP1-amplified GEMM tumors escaping $KRAS^{G12D}$ addiction did not affect IRE1α and proteostasis (fig. S14, A to F). Instead, IRE1α and proteostasis were dependent on ERK and AKT in these cells despite reduced ERK activity compared with that in parental *iKras* cells (fig. S14, G to J). Collectively, these data demonstrate that both reactivated ERK and hyperactivated AKT converge through IRE1α phosphorylation at S525, S529, S549, and T973 to prevent ubiquitination-mediated proteasomal degradation of IRE1α in KRASi-resistant cancer cells. Blocking either individual pathway is not sufficient to inhibit IRE1α because of functional redundancy and compensation by the other pathway.

Next, we sought to understand the mechanism that activates ERK and AKT in the KRASi-resistant cells. Receptor tyrosine kinases (RTKs) activation is one of the most common mechanisms driving sotorasib resistance in patients (*20*–*22*), and RTKs are known to activate ERK and AKT (*17*, *19*, *30*, *80*). Array analysis of 49 RTKs in parental and sotorasib-resistant H358 and MIA-PaCa-2 models revealed the induction of multiple and distinct sets of RTKs in each model (fig. S15, A, B, and E). In the H358 model, EGFR, ErbB2, ErbB3, fibroblast growth factor receptor 3 (FGFR3), and vascular endothelial growth factor receptor 2 (VEGFR2) were substantially induced in the resistant cells (fig. S15, A and B). Inhibiting these RTKs with combined sapitinib, AZD4547, and axitinib, but not individual inhibitor alone, completely suppressed ERK reactivation in $H358^R$ cells (fig. S15, C and D). Blocking FGFR3 with AZD4547 largely abolished AKT hyperactivation (fig. S15, C and D). These up-regulated RTKs had to be simultaneously suppressed to completely blunt both ERK and AKT, resulting in the abrogation of IRE1α restoration in $H358^R$ cells (fig. S15D). Similar to the H358 model, blocking multiple, but not individual, up-regulated RTKs [including EGFR, ErbB2, VEGFR, platelet-derived growth factor receptor β (PDGFRβ), and discoidin domain receptor 2 (DDR2)] completely suppressed both ERK and AKT, leading to the abrogation of IRE1α restoration in a MIA-PaCa-2^R cell (fig. S15, F and G). Treatment of MIA-PaCa-2^R tumors with combined RTK inhibitors—including sapitinib, axitinib, and VU6015929—confirmed the inactivation of ERK and AKT in vivo, leading to the suppression of IRE1α, marked induction of protein aggregation, and reduced tumor growth (fig. S15, H to K). However, they were not well tolerated in the tumor-bearing mice, causing rapid drop of body weight and early lethality (fig. S15L). Collectively, these data demonstrate that multiple and diverse sets of RTKs drive ERK and AKT activation in different KRASi-resistant tumors, which subsequently converge on IRE1α to reestablish proteostasis.

IRE1α inhibition sensitizes oncogenic KRAS-driven tumors to a MEK inhibitor

Although the simultaneous suppression of the MAPK and PI3K pathways, or diverse upstream RTKs, effectively inhibits IRE1α, the heterogeneous resistance mechanisms in different patients (*18*–*23*) and dose-limiting, on-target toxicity of these inhibitors (*81*, *82*) limit their clinical applications for intervening in IRE1α-mediated proteostasis reprogramming in treatment-resistant tumors. Therefore, we directly targeted the IRE1α/XBP1 pathway in KRAS-driven cancers in combination with $KRAS^{G12C}$ or MEK inhibitor. Although MEK inhibitor trametinib or *Xbp1* deletion alone both modestly impeded *iKras* tumor growth in vivo, the loss of *Xbp1* significantly enhanced the response of *iKras* xenograft tumors to trametinib and induced marked protein aggregation (fig. S16, A to D). Treatment of *iKras* tumors with a highly selective IRE1α RNase inhibitor, ORIN1001 (*83*–*86*), recapitulated the effects of the *Xbp1* deletion and markedly enhanced the sensitivity of the *iKras* tumors to trametinib treatment with significant induction of protein aggregation (fig. S16, A and E to G). In a cohort of PDAC patient-derived xenograft (PDX) models (fig. S16N), ORIN1001 also significantly enhanced the sensitivity of the *KRAS*^{G12D}-mutant PATC53 (Fig. 6, A to D), PATC148 (Fig. 6E and fig. S16, H and I), PDAC35 (Fig. 6F and fig. S16, J and K), SW1990 (Fig. 6, G and H, and fig. S16, L and M), and *KRAS*^{G12V}-mutant PDAC19 PDX (Fig. 6I) tumors to trametinib treatment and potently induced protein aggregation in the combination therapy–treated tumors. Collectively, these in vivo data demonstrate that IRE1α/XBP1 inhibition dramatically enhanced the response of *KRAS*-mutant PDAC tumors to trametinib treatment.

IRE1α inhibition enhances the responses of $KRAS^{G12C}$-driven tumors to sotorasib

Next, we examined the effects of IRE1α inhibition on the response to sotorasib treatment in *KRAS*^{G12C}-driven tumors. *IRE1α* silencing modestly reduced MIA-PaCa-2 xenograft tumor growth but not to the extent observed with sotorasib treatment (Fig. 7A and fig. S17A). However, IRE1α deficiency significantly enhanced the response of these tumors to sotorasib treatment and considerably suppressed tumor relapse (Fig. 7A and fig. S17A). By contrast, *PERK* depletion had little impact on MIA-PaCa-2 tumor growth and response to sotorasib, as well as *IRE1α* depletion–induced tumor sensitivity to sotorasib (fig. S17, B to E), excluding the involvement of PERK in KRASi resistance. IRE1α inhibition with ORIN1001 treatment combined with sotorasib treatment also resulted in complete MIA-PaCa-2 tumor regression and long-term remission (Fig. 7B and fig. S17, F and G). We did not observe significant bodyweight changes or signs of toxicity in the combination treatment group (fig. S17, H and I). Treatment of non–KRAS-addicted MIA-PaCa-2^R tumors with ORIN1001 alone also substantially impeded the tumor growth (Fig. 7C) but did not result in complete response. The reduced efficacy with ORIN1001 alone was likely due to the absence of sotorasib, which was required for long-term inhibition of $KRAS^{G12C}$ (*87*), and rewiring of the proteostasis network to be IRE1α-centered.

ORIN1001 has more than 100-fold mammalian enzyme selectivity over its yeast ortholog (*83*). Structure analysis of the mammalian and yeast enzymes revealed a critical residue, Val^{918} (V918), in mammalian IRE1α that differs from the yeast enzyme and could be critical for ORIN1001 binding (Fig. 7D). Binding of ORIN1001

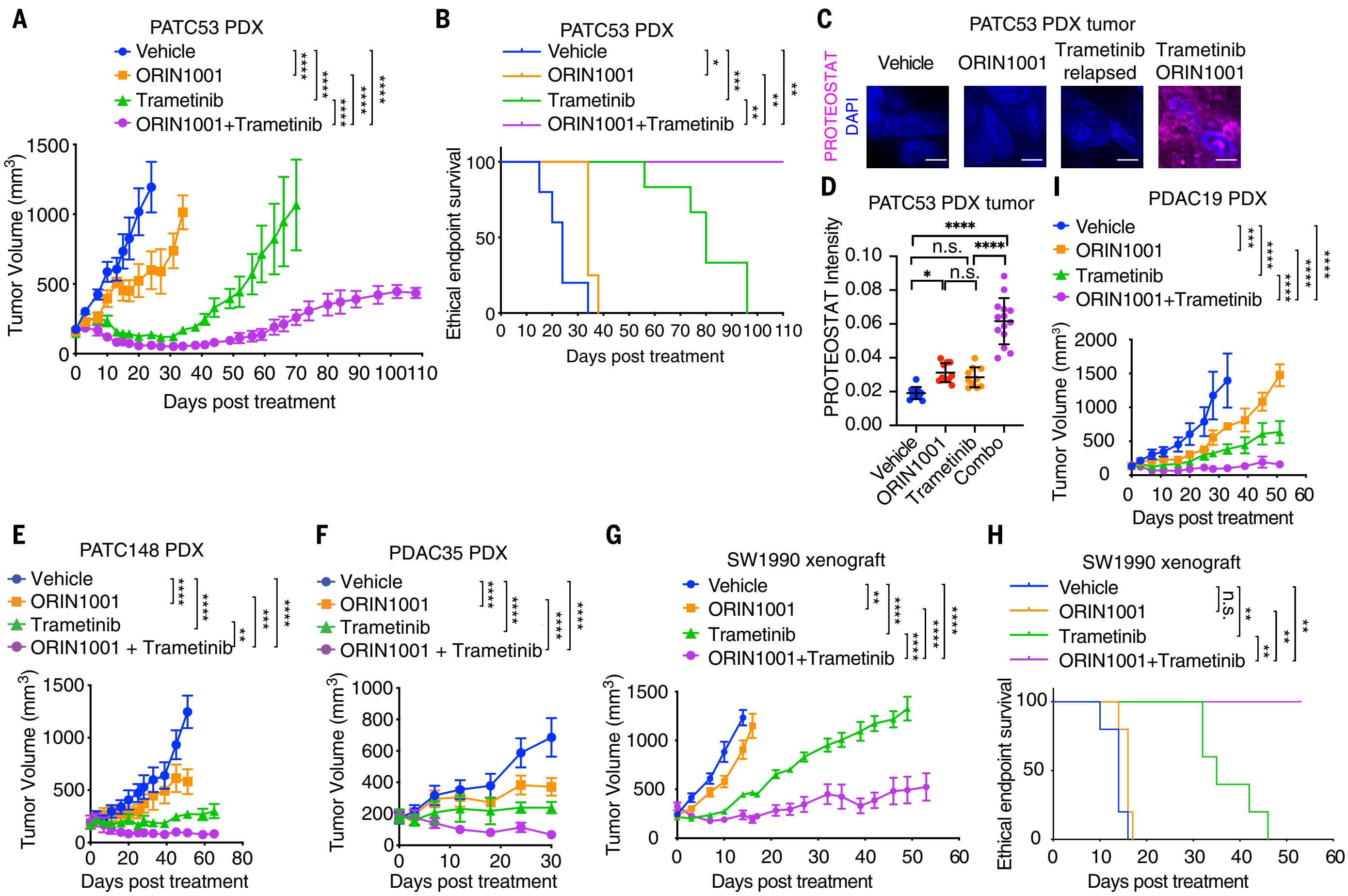


Fig. 6. IRE1α inhibition sensitizes oncogenic KRAS-driven tumors to MEK inhibition. (**A**) Tumor volume quantification of established PATC53 PDX tumors in severe combined immunodeficient (SCID)/beige mice treated with vehicle (n = 5 mice), IRE1α RNase inhibitor ORIN1001 (n = 4 mice), MEK inhibitor trametinib (n = 6 mice), or ORIN1001 plus trametinib (n = 4 mice). (**B**) Kaplan-Meier survival curve of PATC53 PDX tumor-bearing mice under treatment as indicated in (A). (**C**) Representative images and (**D**) quantification of PROTEOSTAT (magenta) and DAPI (blue) staining in endpoint PATC53 xenograft tumors treated as in (A). Data represent average fluorescence intensity of PROTEOSTAT per cell from each image acquired and are presented as mean ± SD from n = 10 independent images. Scale bar, 20 μm. (**E**) Tumor volume quantification of established PATC148 PDX tumors in SCID/beige mice treated with vehicle (n = 6 mice), ORIN1001 (n = 6 mice), trametinib (n = 4 mice), or ORIN1001 plus trametinib (n = 4 mice). (**F**) Tumor volume quantification of established PDAC35 PDX tumors in SCID/beige mice treated with vehicle (n = 5 mice), ORIN1001 (n = 4 mice), MEK inhibitor trametinib (n = 4 mice), or ORIN1001 plus trametinib (n = 4 mice). (**G**) Tumor volume quantification of established SW1990 PDAC xenograft tumors in SCID/beige mice treated with vehicle (n = 6 mice), ORIN1001 (n = 6 mice), trametinib (n = 5 mice), or ORIN1001 plus trametinib (n = 4 mice). (**H**) Kaplan-Meier survival curve of SW1990 PDAC xenograft tumor-bearing mice under treatment as indicated in (G). (**I**) Tumor volume quantification of established PDAC19 PDX tumors in SCID/beige mice treated with vehicle, ORIN1001, trametinib, or ORIN1001 plus trametinib (n = 4 mice). ORIN1001, 150 mg/kg; trametinib, 1 mg/kg. Data are presented as mean ± SEM in (A), (E) to (G), and (I) and mean ± SD in (D). Two-way ANOVA with Bonferroni's multiple comparisons test in (A), (E) to (G), and (I); log-rank (Mantel-Cox) test in (B) and (H); and ordinary one-way ANOVA with Tukey's multiple comparisons test in (D) was used to calculate P values. n.s., not significant; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.

to IRE1α results in the formation of an imine that could be reduced and detected with ultraviolet (UV)–excited fluorescence (Fig. 7E). Whereas purified WT IRE1α protein directly bound to ORIN1001, mutation of valine to phenylalanine at V918 (V918F) abolished the binding (Fig. 7F). The V918F mutation did not affect the ability of ER stressor tunicamycin to induce *XBP1* splicing but completely abolished the response of IRE1α to ORIN1001 (Fig. 7G). Treatment of IRE1α^{V918F}-expressing MIA-PaCa-2^{R} tumors with ORIN1001 failed to inhibit XBP1s in vivo (Fig. 7H). These data identify IRE1α^{V918F} as a drug-resistant mutant that retains intact RNase activity but is immune to ORIN1001. The IRE1α^{V918}-expressing, but not WT IRE1α–expressing, MIA-PaCa-2^{R} tumors were completely immune to ORIN1001-induced sensitivity to sotorasib (Fig. 7I) and protein aggregation (Fig. 7, J and K). Collectively, these data confirm the on-target effects of ORIN1001 in vivo.

We also tested the therapeutic efficacy of combined ORIN1001 and sotorasib treatment in the H358 NSCLC model. The H358 tumors were highly sensitive to sotorasib treatment, and complete regression was observed within 70 days (Fig. 8A). However, the termination of sotorasib treatment resulted in rapid tumor relapse (Fig. 8A). Tumors treated with combined sotorasib and ORIN1001 did not relapse after treatment termination (Fig. 8, A and B). KRASi-resistance mechanisms are heterogeneous in human patients (*18*–*23*). To assess the human relevance, we treated five *KRAS*^{G12C}-mutant NSCLC PDX models with sotorasib and ORIN1001. As shown in Fig. 8, C to G, and fig. S18, A to E, ORIN1001 significantly sensitized all five PDX models to sotorasib. In three PDX models (J000096652, TM00186, and TC303AR), the

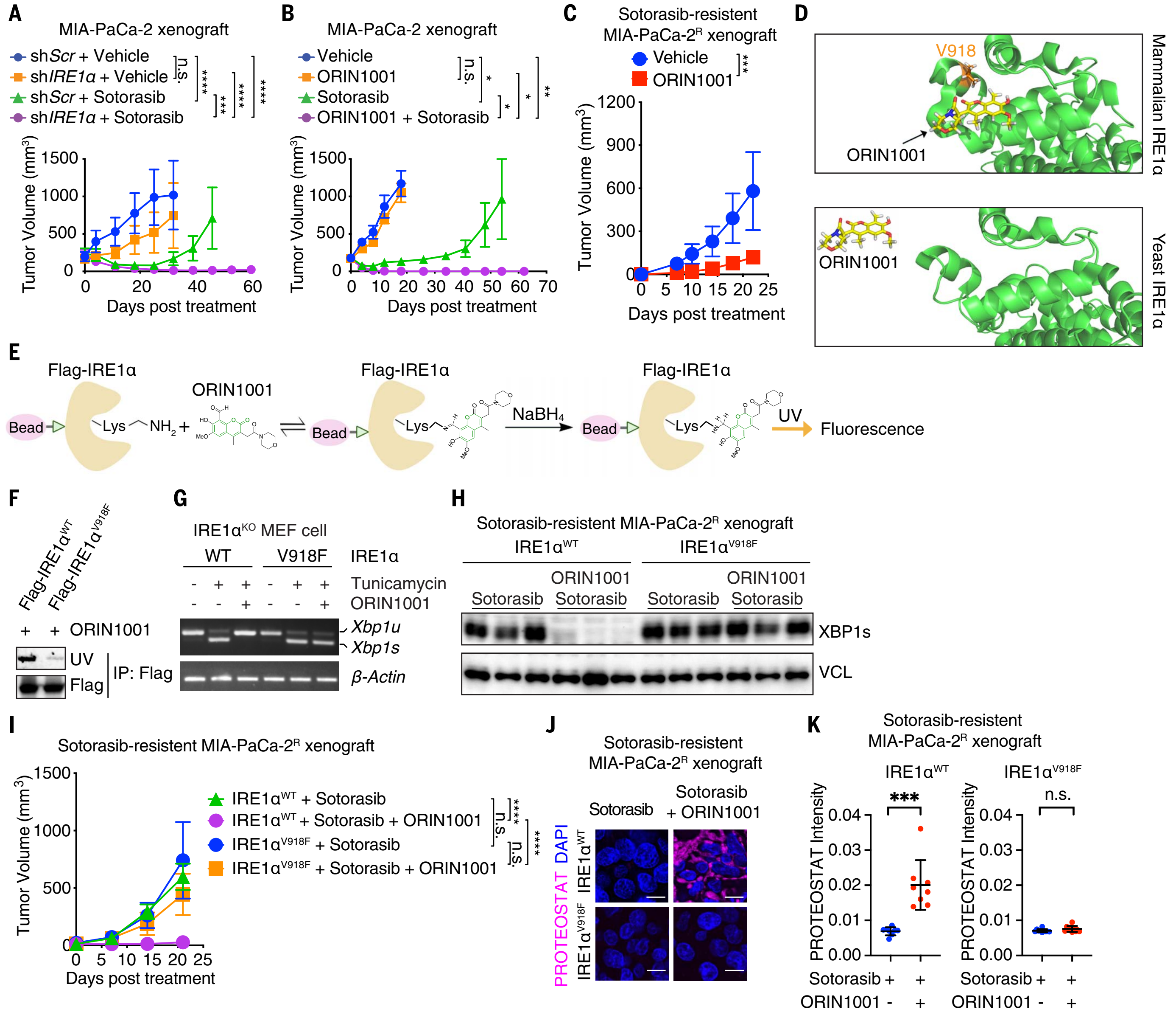


Fig. 7. IRE1α inhibition enhances tumor responses to sotorasib. (**A**) Tumor volume of MIA-PaCa-2 tumors transduced with dox-inducible sh*Scr* or sh*IRE1α* and treated with doxycycline water, vehicle (n = 5 mice) or sotorasib (n = 6 mice). (**B**) Tumor volume of MIA-PaCa-2 tumors treated with vehicle (n = 5 mice), ORIN1001 (n = 4 mice), sotorasib (n = 4 mice), or both (n = 5 mice). (**C**) Tumor volume of sotorasib-resistant MIA-PaCa-2^R tumors treated with vehicle (n = 4 mice) or ORIN1001 (n = 7 mice). (**D**) Docking modeling of ORIN1001 with IRE1α. V918 is critical for the formation of the shallow pocket at mammalian IRE1α RNase-active site for ORIN1001 binding. (**E**) Biochemical fluorescence assay detecting the binding between ORIN1001 and IRE1α in vitro. (**F**) Equal amount of Flag-IRE1α^{WT} or Flag-IRE1α^{V918F} protein purified from 293T cells was used to pull down ORIN1001 in vitro. UV transmission was used to detect ORIN1001 that is covalently bound to IRE1α protein in the SDS–polyacrylamide gel electrophoresis (SDS-PAGE). (**G**) *Xbp1* splicing in *Ire1α*-deleted mouse embryonic fibroblast (MEF) cells expressing IRE1α^{WT} or IRE1α^{V918F} and treated with tunicamycin (5 μg/ml) and/or ORIN1001 (5 μM) for 6 hours as indicated. (**H**) Immunoblot of XBP1s in shRNA-resistant IRE1α^{WT} or IRE1α^{V918F}-transduced, endogenous *IRE1α*-depleted MIA-PaCa-2^R tumors treated as indicated. (**I**) Tumor volume of established shRNA-resistant IRE1α^{WT}- or IRE1α^{V918F}-expressing, endogenous *IRE1α*-depleted, sotorasib-resistant MIA-PaCa-2^R tumors treated as indicated (n = 10 mice). (**J**) Representative images and (**K**) quantification of PROTEOSTAT and DAPI staining in MIA-PaCa-2^R tumors treated as in (I). Data represent average fluorescence intensity of PROTEOSTAT per cell from each image and are presented as mean ± SD from n > 10 independent images. Scale bar, 20 μm. Sotorasib, 100 mg/kg; ORIN1001, 300 mg/kg. Data are presented as mean ± SEM in (A) to (C), and (I). Two-way ANOVA test with Bonferroni's multiple comparisons test in (A) to (C), and (I) and two-tailed, unpaired Student's t test in (K) was used to calculate P values. n.s., not significant; *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

tumors were initially sensitive to sotorasib treatment, but eventually all of the tumors relapsed. Addition of ORIN1001 to sotorasib led to complete responses and prevented tumor relapse (Fig. 8, C to E). In the other two PDX models (J000093018 and TM00192), ORIN1001 also significantly enhanced the tumor responses to sotorasib, but the responses were not as striking as the other models (Fig. 8, F and G). Analysis of these five PDX models showed that sotorasib did not effectively inhibit MAPK in the J000093018 and TM00192 PDX models

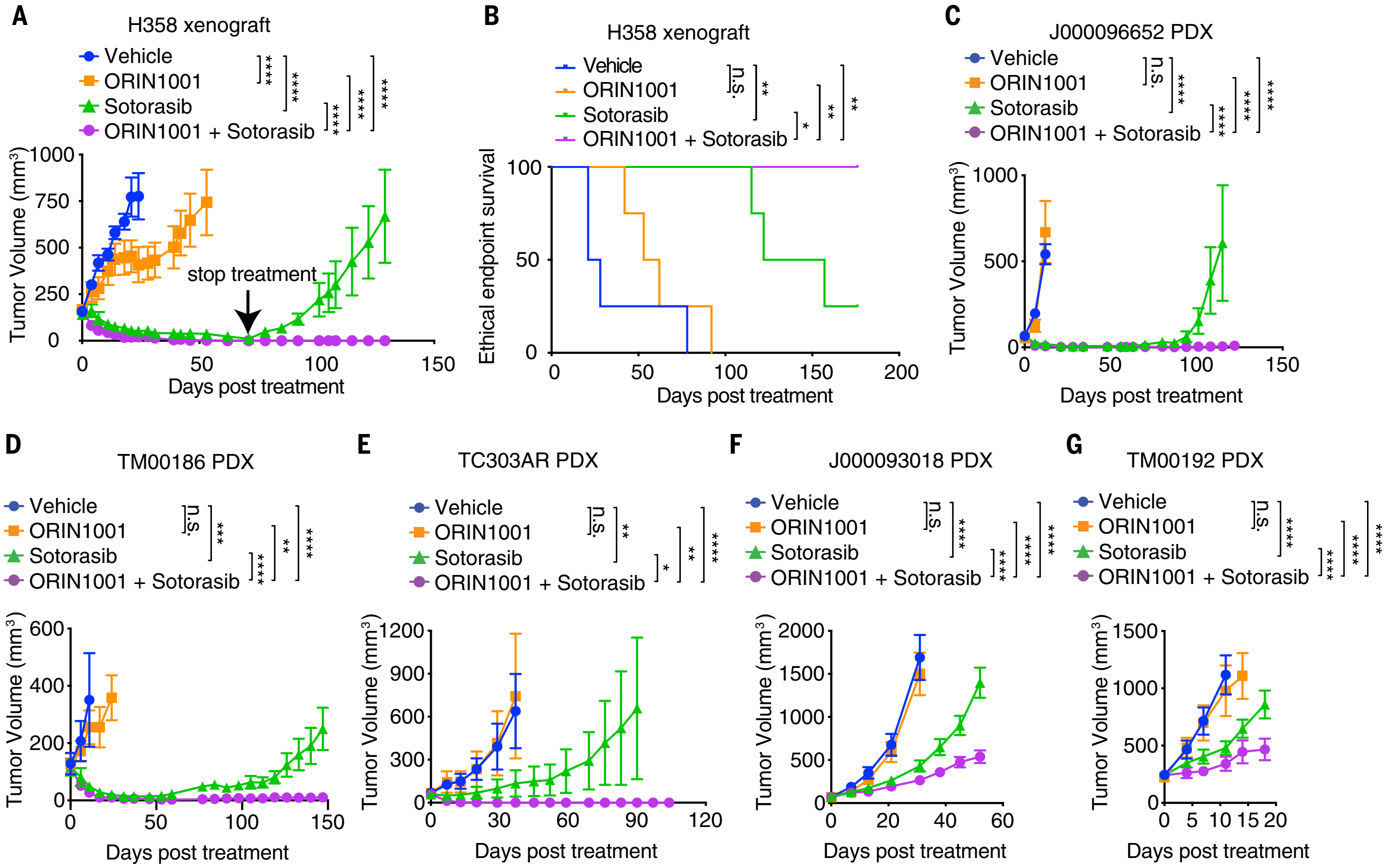


Fig. 8. IRE1α inhibition enhances the response of KRAS^{G12C}-driven tumors to sotorasib. (**A**) Tumor volume quantification of established H358 tumors in SCID/beige mice treated with vehicle, ORIN1001, sotorasib, or ORIN1001 plus sotorasib (*n* = 4 mice). Treatment was stopped at day 71. (**B**) Kaplan-Meier survival curve of H358 tumor-bearing mice under different treatments as indicated in (A) from treatment start time. (**C**) Tumor volume quantification of established J000096652 PDX tumors in NSG mice treated with vehicle (*n* = 6 mice), ORIN1001 (300 mg/kg; *n* = 4 mice), sotorasib (100 mg/kg; *n* = 7 mice), or ORIN1001 plus sotorasib (*n* = 8 mice). Treatment was stopped at day 65. (**D**) Tumor volume quantification of established TM00186 PDX tumors in NSG mice treated with vehicle (*n* = 6 mice), ORIN1001 (300 mg/kg; *n* = 6 mice), sotorasib (100 mg/kg, *n* = 7 mice), or ORIN1001 plus sotorasib (*n* = 9 mice). (**E**) Tumor volume quantification of established TC303AR PDX tumors in NSG mice treated with vehicle (*n* = 5 mice), ORIN1001 (300 mg/kg; *n* = 5 mice), sotorasib (100 mg/kg; *n* = 9 mice), or ORIN1001 plus sotorasib (*n* = 9 mice). Treatment was stopped at day 53. (**F**) Tumor volume quantification of established J000093018 PDX tumors in NSG mice treated with vehicle (*n* = 6 mice), ORIN1001 (300 mg/kg; *n* = 6 mice), sotorasib (100 mg/kg, *n* = 9 mice), or ORIN1001 plus sotorasib (*n* = 9 mice). (**G**) Tumor volume quantification of established TM00192 PDX tumors in NSG mice treated with vehicle (*n* = 6 mice), ORIN1001 (300 mg/kg, *n* = 5 mice), sotorasib (100 mg/kg, *n* = 9 mice), or ORIN1001 plus sotorasib (*n* = 9 mice). Data are presented as mean ± SEM in (A) and (C) to (G). Two-way ANOVA test with Bonferroni's multiple comparisons test in (A) and (C) to (G) and log-rank (Mantel-Cox) test in (B) was used to calculate *P* values. n.s., not significant; **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001.

(fig. S18, F to J), leading to incomplete reprogramming of the proteostasis network and reduced efficacy (fig. S18, K and L). Collectively, these in vivo data demonstrate that IRE1α inhibition is effective in enhancing the response of *KRAS*^{G12C}-driven tumors to sotorasib. Combination therapy with ORIN1001 and sotorasib achieves complete responses and prevents tumor relapse in a significant portion of *KRAS*^{G12C}-driven tumors.

Concluding remarks

Most cancers require a balanced proteostasis network to maintain oncogenic growth. Therapeutic insults often disrupt proteostasis and induce proteotoxic stresses (*88*). How proteostasis network is orchestrated by driver oncogenes and the proteostasis reprogramming mechanisms that bypass oncogene addiction and allow for acquired resistance to targeted therapies remain largely unknown. We have shown that oncogenic KRAS is critical for protein quality control in tumor cells. Inhibition of oncogenic KRAS inactivates both cytosolic and ER protein quality control machinery by inhibiting the master regulators HSF1 and IRE1α. However, residual cancer cells that survive KRAS inhibition directly restore IRE1α through an ER stress–independent unconventional phosphorylation mechanism that reestablishes proteostasis and sustains acquired resistance to KRAS inhibition.

In contrast to what occurs in nonmalignant cells (*53*), oncogenic KRAS activation resolves, rather than induces, ER stress in transformed cancer cells through oncogenic kinase-dependent phosphorylation of IRE1α. We identified four phosphorylation sites in IRE1α that are distinct from IRE1α autophosphorylation sites. The phosphorylation of IRE1α at these sites prevents IRE1α binding with the SEL1L/HRD1 E3 ligase complex, thus impairing the ubiquitination-dependent degradation of IRE1α and stabilizing the protein. These sites are convergence points for multiple resistance pathways and function as a central gatekeeper of the rewired proteostasis network in the KRASi-resistant tumors. Inactivation of these sites is sufficient to abolish the direct regulation of IRE1α by oncogenic signaling and collapse the reestablished proteostasis to overcome resistance to KRASi.

Despite the approval of sotorasib and adagrasib for the treatment of *KRAS*^{G12C}-mutant NSCLC patients, resistance to these inhibitors is rapid and almost inevitable (*15*, *18*–*29*). The heterogeneous resistance mechanisms in patients and dose-limiting toxicity associated with targeting multiple resistance pathways—such as RTKs, MAPK, and PI3K—remain a major barrier to progress. Our mechanistic study directly addressed these clinical challenges by revealing IRE1α-mediated proteostasis reprogramming as a convergence point for multiple heterogeneous resistance mechanisms in response to KRAS-MAPK inhibition. ORIN1001 is a highly specific IRE1α RNase inhibitor (*83*) that demonstrates safety and tolerability in phase I clinical trial (NCT03950570) despite the occurrence of some adverse effects (*86*, *89*). ORIN1001 substantially enhanced the responses of *KRAS*-mutant lung or pancreatic cancer PDX models to sotorasib or trametinib. These data demonstrate that directly targeting IRE1α is a more effective and well-tolerated therapeutic strategy for reversing KRASi or MEKi resistance.

Our study reveals the direct cross-talk between oncogenic signaling and the protein quality control machinery. This study elucidated a molecular mechanism that accounts for the proteostasis modulation observed in response to KRAS inhibition. The mechanisms of KRASi resistance are heterogeneous in patients. Additional studies will be required to examine what proportion of KRAS-driven cancers that develop resistance to KRASi use IRE1α-mediated mechanisms of resistance.

Materials and methods

Cell culture and treatment

MIA-PaCa-2, SW1990, PaTu 8988T, 293T, H358, and BEAS-2B cells were obtained from the American Type Culture Collection (ATCC). Patient-derived PATC53 and PATC148 cells were a gift from Dr. Michael Kim at The University of Texas MD Anderson Cancer Center (*90*). *iKras* cells were derived from our previously generated, doxycycline (Dox)-inducible, *Kras*^{G12D}-driven PDAC mouse model (*tetO_LSL-Kras*^{G12D}*/p53*$^{flox/+}$*/p48-Cre/ROSA26-LSL-rtTA-IRES-GFP*) (*55*). *LSL-Kras*^{G12D} cells were derived from *Kras*^{G12D} knock-in PDAC mouse model (*LSL-Kras*^{G12D}*/p53*$^{flox/+}$) as described previously (*55*). The cell lines used in this study are listed in table S1. BEAS-2B cells were maintained in BEBM Bronchial Epithelial Cell Growth Basal Medium (Lonza, CC-3171) with growth factors and supplements from BEGM Bronchial Epithelial SingleQuots Kit (Lonza, CC-4175). Insulin and hEGF were withdrawn from the medium 48 hours before sample collection. H358, *iKras*, *LSL-Kras*^{G12D}, and PATC148 cells were maintained in RPMI supplemented with 10% FBS serum (Gibco, 10437028) and 100µg/ml penicillin/streptomycin (Invitrogen, 15140163). H358^{R} cells were generated by in vitro culture of H358 cells with increasing dose of sotorasib for 6 months until the cells acquired resistance to 30 nM sotorasib. Doxycycline (VWR, AAJ60579-22, 1µg/ml) was added to the *iKras* cell culture medium to maintain *KRAS*^{G12D} expression. 10% charcoal stripped FBS (VWR, 97065-304) was used to culture *iKras* cells for doxycycline-withdrawal experiments as previously described (*55*). *iKras*^{R} cells were generated by in vitro culture of parental *iKras* cells in the absence of Dox until the cells acquired resistance to *KRAS*^{G12D} inactivation. MIA-PaCa-2, SW1990, PaTu 8988T, 293T and PATC53 cells were maintained in DMEM supplemented with 10% FBS (Gibco, 10437028) and 100µg/ml penicillin and streptomycin (Invitrogen, 15140163). MIA-PaCa-2^{R} cells were generated by in vitro culture of parental MIA-PaCa-2 cells with 30 nM sotorasib until the cells acquired resistance to KRAS inhibition. The chemicals used in this study are listed in table S2.

Tumor inoculation and treatment

The inoculation and establishment of PDX or xenograft tumors was described previously (*90*). PDAC19 and PDAC35 PDX models were generated by Baylor College of Medicine PDAC PDX Core. TC303AR PDX was generated by MD Anderson Cancer Center PDX Core. J000096652, TM00186, J00093018, and TM00192 PDX models were purchased from Jackson Laboratory. For tumor fragments transplantation of PDX, 1mm^3 fresh tumor fragments were transplanted into the lower flanks of 6-week-old immune-compromised SCID/Beige mice (Charles river, strain code 250, CB17.Cg-*Prkdc*^{scid}*Lyst*$^{bg\text{-}J}$*/Crl*, both female and male) or NSG mice (Jackson Laboratory, stain code 005557, NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}*/SzJ*, both female and male). For subcutaneous xenograft experiments, 1×10^6 PATC53, SW1990, MIA-PaCa-2, H358 or *iKras* cells suspended in 100 µl 50% Matrigel (Corning, #354230, in PBS) were injected subcutaneously into the lower flanks of 6-week-old SCID/Beige mice or Athymic Nude mice (Envigo, strain code 69, Hsd:Athymic Nude-*Foxn1*^{nu}, female). Tumor growth was monitored using calipers and tumor volumes were calculated by the equation $V(\text{mm}^3) = L \times W^2/2$, where L is the largest diameter and W is the perpendicular diameter. When tumors reached a volume of approximately 50–500 mm^3, mice were randomized and treated with drugs as indicated. ORIN1001 was provided by Orinove and suspended in 1% microcrystalline cellulose in 50% sucrose and sonicated for 90 min in water bath sonicator (VWR Ultrasonic Cleaner, Model 97043-964) before dosing at 150 mg/kg or 300 mg/kg body weight via daily oral gavage (*83*). Trametinib (1mg/kg) or sotorasib (30, 50, or 100 mg/kg) was formulated in the hydroxypropylmethylcellulose (HPMC)-Tween 80 buffer solution (0.5% HPMC and 0.2% Tween 80, pH 8.0) and administered via daily oral gavage as described previously (*12*). SCH772984 (50mg/kg) was formulated in 45% saline, 50% PEG 400 and 5% DMSO, and administered via daily intraperitoneal injection. MK2206 (120 mg/kg) was formulated in 30% Captisol (Cydex) and administered by oral gavage every other day. All mice were maintained in accordance with Baylor College of Medicine Animal Care and Use Committee procedures and guidelines.

Protein aggregation detection assay

The PROTEOSTAT Aggresome detection kit (Enzo Life Sciences, ENZ-51035-K100), Congo red dye (CR, Sigma, 234610) or thioflavin T dye (ThT, Sigma, T3516) was used to detect misfolded or aggregated proteins in cells or tumor tissues (*56*, *72*). The PROTEOSTAT aggresome detection assay was performed according to the manufacturer's instructions. Briefly, the *iKras* cells seeded on glass slides were washed with PBS, fixed with 4% formaldehyde for 30 min at room temperature (RT), permeabilized with Permeabilizing Solution (0.5% Triton X-100, 3mM EDTA) for 30 min on ice with gentle shake, and stained with the PROTEOSTAT dye (1:20,000 dilution) for 30 min at RT. MIA-PaCa-2 or H358 cells were trypsinized from culture dishes followed by washing, fixation, permeabilization and staining as described above. Tumor sections were deparaffinized and rehydrated before staining. Samples were incubated with PROTEOSTAT dye (1:20,000 dilution) for 30 min at RT. Nuclei were counterstained with DAPI. Cells treated with 10 µM MG132 (provided in the PROTEOSTAT Aggresome detection kit) for 16 hours was used as positive control. Samples stained with DAPI only were used as negative control. Congo red (CR) and thioflavin T (ThT) staining were performed as described previously (*56*). Briefly, the *iKras* cells seeded on glass slides were washed with PBS, fixed with 4% formaldehyde for 30 min at RT, permeabilized with Permeabilizing Solution (0.5% Triton X-100, 3mM EDTA) for 30 min on ice with gentle shake, and stained with 20 µM ThT or 50 nM CR dissolved in PBS for 30min at RT, followed by rinsing in PBS for 3 times. Nuclei were stained with DAPI.

Images (16-bit greyscale TIFFs) were analyzed using CellProfiler v2.2 (Broad Institute) as described previously (*72*). Briefly, the DAPI channel images were first smoothened with a median filter and nuclei were identified with automatic thresholding and a fixed diameter. The cell nuclei that touch the border of the image were eliminated for quantification. The cell nuclei that touch each other were separated with a watershed algorithm. Then, cell boundaries were identified by watershed gradient based on the dye signal of PROTEOSTAT, ThT or CR, using nuclei as a seed. Metrics for PROTEOSTAT, ThT, or CR were extracted from the cells.

Cell viability assay

Two hundred cells were seeded in 96-well plate and treated with different inhibitors as indicated in the figures. Cell viability was measured daily with CCK-8 kit (APExBIO, # K1018). Briefly, 10 µl CCK-8 solution was added to each well and incubated for 1 hour at 37°C. The absorbance at 450 nm was measured using BioTek Synergy HTX Multi-Mode microplate reader.

Colony formation assay

Two hundred cells were seeded in 12-well plate and cultured for 5 days. Cells were washed with PBS, fixed with methanol and stained with 0.5% crystal violet. The colony number was counted and quantified.

BrdU incorporation assay

For BrdU staining, cells in the logarithmic phase of proliferation were first labeled with BrdU at a final concentration of 10 µM for 30 min at 37°C, followed by intracellular staining using the BrdU staining kit according to the manufacturer's instructions (BD Biosciences, 559619). Flow cytometry data were collected using BD FACS Diva 8 on a BD LSR II or BD Fortessa analyzer. The acquired data were analyzed using the FlowJo 10 software.

Luciferase assay

For *iKras* cells, the HSF1 firefly luciferase reporter was constructed by cloning 4 copies of the heat shock element (HSE) followed by a minimal promoter sequence (5′-AGAGGGTATATAATGGAAGCTCGACTTCCAG-3′) (Promega, E375A) (all primer sequences are listed in table S3) into the SacI and HindIII sites of the pGL3-basic luciferase reporter (Promega, E1751). The construct was verified by DNA sequencing. The $iKras^P$ or $iKras^R$ cells were seeded in 96-well plate at 800 cells/well and co-transfected with 100ng *firefly* luciferase reporter and 5ng *Renilla* luciferase plasmid (pRL-TK, Promega, E2241, used as internal control) using Lipofectamine 3000 (Thermo Fisher, L300008). The transfected $iKras^P$ cells cultured in the presence or absence of Dox for 48 hours or the transfected $iKras^R$ cells cultured in the absence of Dox were heat shocked at 43°C for 1 hour and recovered overnight before measuring the luciferase activities using the Dual-luciferase Reporter Assay System (Promega, #E1910) according to the manufacturer's instructions. For MIA-PaCa-2 cells, the same sequence (Promega, E375A) was cloned into the XhoI and BamHI sites of the pRRL-Luciferase plasmid (Addgene, 120798). The construct was verified by DNA sequencing. The *firefly* or pLenti.PGK.blast-Renilla_Luciferase (Addgene, 74444) plasmid was packaged into lenti-viruses and infected MIA-PaCa-2^P or MIA-PaCa-2^R cells. After selection with blasticidine (20 µg/ml), the cells were seeded in 96-well plate at 800 cells/well. The infected MIA-PaCa-2^P cells cultured in the presence or absence of sotorasib for 48 hours were heat shocked at 43°C for 1 hour and recovered overnight before measuring the luciferase activities using the Dual-luciferase Reporter Assay System (Promega, #E1910) according to the manufacturer's instructions.

Proteasome activity assay

The Proteasome Activity Fluorometric Assay Kit (BioVision, K245) was used to detect proteasome activity according to the manufacturer's instructions. Briefly, $1x10^6$*iKras* cells from different treatments were homogenized in a tight-fitting bounce homogenizer (Thomas Scientific, 1176F27) with 500µl 0.5% NP-40 in PBS. 10 µl Proteasome Substrate (AMC-peptide, provided in the kit) was added to 10 µl cell lysate from each treatment group or 10 µl positive control lysate (provided in the kit) and mixed. The reaction was performed at 37°C and protected from light. The kinetics of fluorescence at excitation/emission = 350/440 nm were measured every 30 min using BioTek Synergy HTX Multi-Mode Plate Reader. Cells treated with 10µM MG132 (provided in the detection kit) for 16 hours were used as negative control. The fluorescence signals were normalized against total protein abundance detected with BCA Protein Assay Kit (Thermo Fisher Scientific, 23225).

Plasmids, virus production, and infection

The plasmids used are listed in table S4. The pRK5-Flag-IRE1α or pCDH-Flag-IRE1α was generated by cloning the full-length human IRE1α into pRK5 (Genentech) or pCDH (System Biosciences, CD511B-1) vector. The point mutations of IRE1α were introduced with Q5 Site-Directed Mutagenesis Kit (New England Biolabs, E0554). All the IRE1α plasmids are shRNA-resistant and listed in table S4. Primers used for cloning are listed in table S3. pHAGE-$BRAF^{V600E}$, pHAGE-MEK^{DD}, and pHAGE-$PIK3CA^{H1047R}$ plasmids were generated as described previously (*91*). The pCDH-HA-Myr-AKT plasmid was purchased from Addgene (# 46969) (*92*). pLVX-Flag-HA-*HRD1* was generated as described previously (*73*). The shRNAs targeting mouse *Xbp1*, *Ire1α*, *Perk*, *Gcn2*, *Hri*, *Pkr*, and *Scp3* were cloned into pLKO.1-TRC (Addgene, 10878). The shRNAs targeting human *XBP1*, *IRE1α, NcK,* and *PERK* were cloned in pLKO.1-TRC (Addgene 10878) or pLKO-Tet-On (Addgene 21915) vector to generate constitutive or inducible constructs. The shRNA sequences are listed in table S3. The pLKO.1 shScramble (Addgene 1864) or pLKO.1 Tet-On shScramble (same shRNA sequence as in pLKO.1 shScramble) was used as control. The p-GIPZ non-silencing shRNA control, p-GIPZ-MAPK1, p-GIPZ-MAPK3 were from Dharmacon Reagents. The *Cas9*-expressing plasmid lentiCas9-Blast was purchased from Addgene (#52962) (*93*). The gRNAs targeting *Ire1α* or *Xbp1* were cloned into lentiGuide-Puro vector (Addgene, #52963). All gRNA sequences are listed in table S3. Plasmids containing coding sequence of different phosphatases are listed in table S4. To generate lentiviruses, 293T cells were co-transfected with psPAX2 and pMD2.G using Lipofectamine 3000 (Thermo Fisher Scientific, # L3000008). Lentiviruses were collected 48 and 72 hours after transfection and used for infecting cells in the presence of 8 µg/ml polybrene (Millipore Sigma, TR-1003-G) prior to puromycin selection (2 µg/ml, Millipore Sigma, P8833).

Generation of knock-out (KO) cells

To generate *Ire1α* or *Xbp1* KO cells, *iKras* cells were first infected with lentiviruses encoding *Cas9* (lentiCas9-Blast, Addgene 52962) (*93*) and selected with 10 µg/ml blasticidin (Santa Cruz, 3513-03-09). The *Cas9*-expressing cells were then infected with lentiviruses expressing two gRNAs targeting the same exon and selected with puromycin (2µg/ml, Millipore-Sigma, P8833) to generate pooled KO cells. The gRNA sequences are listed in table S3.

Immunohistochemical (IHC) staining

Tumor specimens were fixed with freshly made 4% paraformaldehyde for 24 hours, washed with PBS and stored in 70% ethanol until paraffin embedding. IHC staining was performed on 5 µm-thick paraffin sections. For p-ERK, p-HSF1, p-GCN2, p-eIF2α and YAP1 staining, 10mM sodium citrate buffer (pH 6.0) was used for antigen retrieval. For IRE1α, p-PERK, ATF4 and p-AKT IHC, 1mM EDTA buffer (pH 9.0) was used for antigen retrieval. Endogenous peroxidase was quenched with 3% H_2O_2 for 20 min followed by blocking with 3% normal goat serum. The following primary antibodies were used: IRE1α (1:20, Cell Signaling Technology, 3294); p-ERK (1:200, Cell Signaling Technology, 4376); and p-HSF1 (1:200, Life Technologies, BSM-52166R); p-PERK (1:25, Cell Signaling Technology, 3179); ATF4 (1:50, Santa Cruz, 390063); p-GCN2 (1:200, Thermo Fisher, PA5-105886); p-AKT (1:50, Cell Signaling Technology, 4060); p-eIF2α(1:50, Cell Signaling Technology, 9721); YAP1 (1:400, Cell Signaling Technology, 14074). Slides were incubated with ImmPRESS Excel HRP Goat Anti-Rabbit Polymer Reagent (Vector labs, MP-7451-15) for 30 min. Sections were developed with DAB+ solution (Dako, K3468) and counterstained with Harris Hematoxylin. The antibody used are listed in table S5.

Tissue microarray (TMA) analysis

For quantifications of TMA staining, TMAs stained with anti-IRE1α or anti-p-ERK antibody were scanned using the Aperio scanner and analyzed with QuPath software (*94*). Detailed tutorials of the software can be found at https://qupath.readthedocs.io/en/stable/index.html.

Briefly, images were preprocessed by automated “TMA dearraying” and “stain” vector estimation. Tissue sections were identified by running “simple tissue detection.” The “positive cell detection” command was used to detect DAB staining intensity. The score compartment was set as “DAB OD mean.” Tumor cells and stromal cells were classified by “training object classifier” based on annotations. The fraction score was calculated as the proportion of positively stained tumor cells (0%-100%). The intensity and fraction scores were then multiplied to obtain the H-score.

RNA extraction and real-time quantitative reverse-transcriptase PCR

Total RNA was extracted using TRIzol (Thermo Fisher Scientific, 15596026). Total RNA (1 μg) was converted to cDNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, 4368813), followed by qPCR on a QuantStudio 6 Flex Real-Time PCR System (Applied Biosystems). The sequences of all primers are listed in table S3.

Detergent-insoluble aggregates detection

Detergent-insoluble aggregates detection was performed as described previously (*47*). Briefly, cells with different treatments were harvested by trypsinization. After washing with cold PBS, 1×10^6 cells were lysed with RIPA buffer (25 mM Tris-HCl pH7.6, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1.5 mM EDTA) supplemented with protease inhibitor cocktail (Roche, 14826500), phosphatase inhibitor cocktail (Sigma, 4906845001) and 10 mM N-Ethylmaleimide (Sigma, E3876). Protein lysates were cleared by centrifugation at 16,000 g for 20 min at 4°C. The remaining insoluble pellets were then washed with RIPA buffer for three times to remove any remaining detergent-soluble proteins. The pellet containing detergent-insoluble aggregates was solubilized with urea buffer (8 M urea, 2% SDS, 50 mM DTT, 50 mM Tris-HCl pH7.4) for Western blot analysis. 3×10^5 cells were directly lysed in urea buffer (8 M urea, 2% SDS, 50 mM DTT, 50 mM Tris-HCl pH7.4) serving as loading control.

Coimmunoprecipitation

The co-immunoprecipitation (co-IP) assay was performed as previously described (*73*). Briefly, 293T cells transfected with different plasmids or H358 cells infected with indicated viruses were lysed with lysis buffer (50mM Tris-HCl pH7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100) supplemented with protease inhibitor cocktail (Roche, 14826500) and phosphatase inhibitor cocktail (Sigma, 4906845001). Protein lysates were cleared by centrifugation at 12,000 g for 20 min at 4°C. Supernatant was incubated with anti-Flag M2 beads (Sigma, F-2426) or anti-Myc beads (Sigma, E-6654) for 4h to overnight at 4°C with gentle rotating. The beads were then washed once with lysis buffer, followed by additional three washes with wash buffer (50mM Tris-HCl pH7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 10% glycerol). The proteins were eluted by boiling in 2 × Laemmli sample buffer [65.8 mM Tris-HCl, pH 6.8, 2.1% SDS, 26.3% (w/v) glycerol, 355 mM β-mercaptoethanol, 0.01% bromophenol blue] and analyzed by SDS-PAGE and Western blot.

Ubiquitination and phosphorylation assay

The ubiquitination and phosphorylation assays were performed to detect IRE1α ubiquitination and phosphorylation. MIA-PaCa-2 or H358 cells infected with lentiviruses encoding control or Flag-IRE1α were treated with DMSO, sotorasib or trametinib as described in figure legend and lysed with RIPA buffer (25 mM Tris-HCl pH7.6, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1.5 mM EDTA) supplemented with protease inhibitor cocktail (Roche, 14826500), phosphatase inhibitor cocktail (Sigma, 4906845001) and 10 mM N-Ethylmaleimide (Sigma, E3876). Protein lysates were sonicated for 30s and cleared by centrifugation at 12,000 g for 20 min at 4°C. Supernatant was incubated with anti-Flag M2 beads (Sigma, F-2426) for 4h at 4°C with gentle rotating. The beads were then washed with RIPA buffer for three times. The immunoprecipitates were eluted and denatured by boiling for 10 min in denature buffer (50 mM Tris-HCl pH7.6, 1% SDS, 0.5 mM EDTA, 1 mM DTT) to disrupt the interactions between immunoprecipitated IRE1α and its interacting proteins. The denatured elutes were diluted 1:10 with lysis buffer (50mM Tris-HCl pH7.6, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100) and subjected to a second-round immunoprecipitation with anti-Flag M2 beads (Sigma, F-2426) (4h at 4°C) to remove all the interacting proteins and selectively pull down only IRE1α protein which allows for the specific analysis IRE1α ubiquitination and phosphorylation. The beads were then washed with RIPA buffer for three times. The proteins were eluted by boiling in 2 × Laemmli sample buffer [65.8 mM Tris-HCl pH 6.8, 2.1% SDS, 26.3% (w/v) glycerol, 355 mM β-mercaptoethanol, 0.01% bromophenol blue] and analyzed by SDS-PAGE and Western blot.

GST pull-down assay

The GST pull-down assay was performed to detect interaction between ERK and IRE1α. Briefly, recombinant GST or GST-ERK2 protein purified from *E.coli* (SignalChem, M28-10G-20) was incubated with recombinant His-IRE1α protein purified from Sf9 cells (*83*) in RIPA buffer (25 mM Tris-HCl pH7.6, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1.5 mM EDTA) for 30 min before the GSH-Sepharose beads (GE Healthcare, 17075601) were added and rotation for 1 hour at 4°C . The beads were washed with RIPA buffer for three times. The proteins were eluted by boiling in 2 × Laemmli sample buffer and analyzed by SDS-PAGE and Western blot.

Flag pull-down assay

293T cells transfected with plasmid expressing Flag-GFP or Flag-IRE1α were lysed in RIPA buffer (25 mM Tris-HCl pH7.6, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1.5 mM EDTA, supplemented with protease inhibitor cocktail and phosphatase inhibitor cocktail) 48 hours after transfection. Protein lysates were cleared by centrifugation at 12,000 g for 20 min at 4°C. Supernatant was incubated with anti-Flag M2 beads (Sigma, F-2426) overnight at 4°C with gentle rotating. The beads were then washed with RIPA buffer for three times. These preloaded Flag M2 beads with Flag-GFP or Flag-IRE1α were then incubated with purified GST-ERK2 (SignalChem, M28-10G-20) for 30 min before washing with wash buffer (50mM Tris-HCl pH7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100) for three times and elution by boiling in 2 × Laemmli sample buffer for 10 min. The eluents were analyzed by SDS-PAGE and Western blot.

In vitro kinase assay

Flag-GFP or Flag-IRE1α proteins were purified from 293T cells as described above. The Flag M2 beads pre-loaded with Flag-GFP or Flag-IRE1α proteins were rinsed with kinase assay buffer I (SignalChem, K01-09, 25 mM MOPS pH7.2, 12.5 mM β-glycerol-phosphate, 25 mM $MgCl_2$, 5 mM EGTA, 2 mM EDTA, 0.25 mM DTT) and subjected to in vitro kinase assay in the kinase assay buffer I plus 0.4 mM cold ATP and 1.0 μg GST-ERK2 activated by MEK1 in vitro (SignalChem, M28-10G-20). The reaction was carried out at 30°C for 30 min. The beads were then washed for three times with RIPA buffer (25 mM Tris-HCl pH7.6, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1.5 mM EDTA). The proteins were eluted by boiling in 2 × Laemmli sample buffer and analyzed by SDS-PAGE and Western blot.

For in vitro kinase assay labeled with [γ-^{32}P] ATP, Flag-IRE1α^{K599A} (kinase-dead form) or different mutated Flag-IRE1α^{K599A} proteins expressed in 293T cells were bound to anti-Flag M2 beads as described above and incubated with 0.5 μg GST-ERK2 activated by MEK1 in vitro (SignalChem, M28-10G-20) in the presence of 1 μCi of [γ-^{32}P] ATP (PerkinElmer, NEG002A100UC) and 0.4 mM cold ATP in the kinase buffer I (SignalChem, K01-09) for 30 min at 30°C. The reaction was stopped by boiling in 2 × Laemmli sample buffer for 10 min. The eluted proteins were resolved by SDS-PAGE and detected by autoradiography.

For in vitro kinase assay with recombinant His-IRE1α, 1.0 μg protein purified from Sf9 cells (*83*) and recombinant GST-ERK2 protein,

the reaction was carried out in the kinase assay buffer I (SignalChem, K01-09) plus 0.4 mM cold ATP for 30 min at 30°C. The reaction was stopped by adding 8 M urea buffer followed by purification of His-IRE1α with Ni-NTA agarose (QIAGEN, 30210). The proteins were eluted with 2 × Laemmli sample buffer by boiling for 10 min and analyzed by SDS-PAGE and Western blot.

In vitro phosphatase assay

The in vitro phosphatase assay was performed as described previously (*95*). Phosphorylated Flag-IRE1α proteins were purified from 293T cells expressing MEK^{DD} as described above. The Flag M2 beads pre-loaded with phosphorylated Flag-IRE1α proteins were rinsed with phosphatase assay buffer (40 mM Tris-HCl pH7.5, 20 mM KCl, 10 mM $MnCl_2$, and 2 mM DTT) and subjected to in vitro phosphatase assay in the phosphatase assay buffer and 1.0 μg recombinant His-SCP3 protein (NOVUS, NBP1-99109). The reaction was carried out at 30°C for 30 min. The beads were then washed for three times with RIPA buffer (25 mM Tris-HCl pH7.6, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1.5 mM EDTA). The proteins were eluted by boiling in 2 × Laemmli sample buffer and analyzed by SDS-PAGE and Western blot. For in vitro kinase assay with recombinant His-IRE1α, 1.0 μg His-IRE1α protein was phosphorylated by recombinant GST-ERK2 protein in vitro as described above, and then subjected to in vitro phosphatase assay with SCP3 protein. 1.0 μg recombinant His-SCP3 protein (NOVUS, NBP1-99109) or Myc-SCP3 protein purified from H358 cells treated with DMSO or sotorasib (30 nM) were used for the in vitro phosphatase assay as indicated in the figures. The proteins were boiled in 2 × Laemmli sample buffer and analyzed by SDS-PAGE and Western blot.

Phospho-RTK array

The Proteome Profiler Human Phospho-RTK Array Kit (R&D systems, ARY001B) was used to assess the phosphorylation status of 49 RTKs in H358 cells or MIA-PaCa-2 tumors treated with vehicle or sotorasib. The phospho-RTK array assay was performed according to the manufacturer's instructions. Briefly, the H358 cells or MIA-PaCa-2 tumors were lysed in the Lysis Buffer 17 (provided in the kit) with protease inhibitors. The protein concentration was measured with Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific, 23225). 1000 μg total protein lysate was incubated with the RTK array membrane overnight at 4°C. After washing, the anti-phospho-tyrosin-HRP detection antibody was incubated with the membrane, and chemiluminescence signal was detected with Amersham Imager 600 (GE Healthcare Lifer Sciences) and quantified with Image J.

Biochemical fluorescence assay to detect IRE1α interaction with ORIN1001

The in vitro binding assay to detect IRE1α with ORIN1001 was performed as described previously (*96*). Briefly, WT or V918F mutant Flag-IRE1α proteins were purified from 293T cells as described above. The Flag M2 beads pre-loaded with equal amount of Flag-$IRE1\alpha^{WT}$ or Flag-$IRE1\alpha^{V918F}$ proteins were incubated with 200 μM ORIN1001 in binding buffer (50 mM Tris-HCl pH7.5, 100 mM NaCl, 10% Glycerol, 1% Triton X-100, 50 mM EDTA) for 3 hours at 4°C. After that, 6 mM $NaBH_4$ was added to reduce the imine (Schiff base) between IRE1α and ORIN1001 to stable amine. The beads were then extensively washed for three times with binding buffer and three times with RIPA buffer (25 mM Tris-HCl pH7.6, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1.5 mM EDTA). The proteins were eluted by boiling in 2 × Laemmli sample buffer and analyzed by SDS-PAGE. UV transmission (GE Healthcare Lifer Sciences) was used to detect ORIN1001 that is covalently bound to IRE1α protein in the SDS-PAGE.

Modelling of the IRE1α-ORIN1001 complex

Murine IRE1α and MKC9989 (PDB 4PL3) was used as a template to model ORIN1001 bound to murine IRE1α with Schrodinger software. Homology modelling of yeast IRE1α (PDB 3FBV) was also achieved and was superimposed manually onto the model of ORIN1001 with murine IRE1α.

Clinical Proteomic Tumor Analysis Consortium (CPTAC) data analysis

Correlation between IRE1α phosphorylation (S549) and ERK phosphorylation (Y204) in CPTAC NSCLC tumors were analyzed using the LinkedOmics platform (*97*) (http://www.linkedomics.org).

Western blot

Whole cell lysates or immunoprecipitation samples were separated by SDS-PAGE and transferred to nitrocellulose membrane (Biorad, 1620112). The antibody used are listed in table S5. The reagents and kits used in this study are listed in table S2.

Statistics and reproducibility

Data are expressed as the mean ± SD or mean ± SEM as indicated in the figure legends; n is the number of independent biological replicates, unless specifically indicated otherwise in the figure legend. The respective n values are shown in the figure legends. No statistical method was used to pre-determine the sample sizes. For animal experiments, at least 4 biological replicates were included based on previously published work, preliminary studies as standard for this field of research. See figures legends for each experiment. Treatments were performed in a non-blinded manner by a research technician who was not aware of the objectives of the study because the colors of the drugs used are different from that of vehicle control. The results were quantified using GraphPad Prism 8. Two-tailed, unpaired Student's t test with or without Welch's correction was utilized to compare the differences between 2 groups as indicated in figure legends. One-way ANOVA with Dunnett's or Tukey's multiple comparison test was used to compare the differences among 3 or more groups as indicated in figure legends. Two-way ANOVA with Bonferroni's multiple comparison test was used to calculate the significance difference for cell growth, tumor volume and body weight measurement over time. The log-rank (Mantel-Cox) test was used to test for the significant differences of survival between the groups. $P < 0.05$ was considered statistically significant. No samples or animals were excluded from the analysis.

Study approval

All protocols described in this study were approved by the Baylor College of Medicine Institutional Animal Care and Use Committee (AN6813).

REFERENCES AND NOTES

1. A. J. Scheidig, C. Burmester, R. S. Goody, The pre-hydrolysis state of $p21^{ras}$ in complex with GTP: New insights into the role of water molecules in the GTP hydrolysis reaction of ras-like proteins. *Structure* **7**, 1311–1324 (1999). doi: 10.1016/S0969-2126(00)80021-0; pmid: 10574788
2. K. Scheffzek *et al*., The Ras-RasGAP complex: Structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* **277**, 333–338 (1997). doi: 10.1126/science.277.5324.333; pmid: 9219684
3. G. Bollag, F. McCormick, Differential regulation of rasGAP and neurofibromatosis gene product activities. *Nature* **351**, 576–579 (1991). doi: 10.1038/351576a0; pmid: 1904555
4. C. Li *et al*., The G protein signaling regulator RGS3 enhances the GTPase activity of KRAS. *Science* **374**, 197–201 (2021). doi: 10.1126/science.abf1730; pmid: 34618566
5. J. P. McGrath, D. J. Capon, D. V. Goeddel, A. D. Levinson, Comparative biochemical properties of normal and activated human ras p21 protein. *Nature* **310**, 644–649 (1984). doi: 10.1038/310644a0; pmid: 6147754
6. J. Downward, Targeting RAS signalling pathways in cancer therapy. *Nat. Rev. Cancer* **3**, 11–22 (2003). doi: 10.1038/nrc969; pmid: 12509763
7. D. K. Simanshu, D. V. Nissley, F. McCormick, RAS proteins and their regulators in human disease. *Cell* **170**, 17–33 (2017). doi: 10.1016/j.cell.2017.06.009; pmid: 28666118
8. A. D. Cox, S. W. Fesik, A. C. Kimmelman, J. Luo, C. J. Der, Drugging the undruggable RAS: Mission possible? *Nat. Rev. Drug Discov.* **13**, 828–851 (2014). doi: 10.1038/nrd4389; pmid: 25323927
9. J. Encarnación-Rosado, A. C. Kimmelman, Harnessing metabolic dependencies in pancreatic cancers. *Nat. Rev. Gastroenterol. Hepatol.* **18**, 482–492 (2021). doi: 10.1038/s41575-021-00431-7; pmid: 33742165
10. Y. Pylayeva-Gupta, E. Grabocka, D. Bar-Sagi, RAS oncogenes: Weaving a tumorigenic web. *Nat. Rev. Cancer* **11**, 761–774 (2011). doi: 10.1038/nrc3106; pmid: 21993244
11. J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells, K. M. Shokat, K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013). doi: 10.1038/nature12796; pmid: 24256730
12. J. Canon *et al*., The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature* **575**, 217–223 (2019). doi: 10.1038/s41586-019-1694-1; pmid: 31666701
13. J. Hallin *et al*., The $KRAS^{G12C}$ inhibitor MRTX849 provides insight toward therapeutic susceptibility of KRAS-mutant cancers in mouse models and patients. *Cancer Discov.* **10**,

54–71 (2020). doi: 10.1158/2159-8290.CD-19-1167; pmid: 31658955
14. F. Skoulidis *et al*., Sotorasib for lung cancers with *KRAS* p.G12C mutation. *N. Engl. J. Med.* **384**, 2371–2381 (2021). doi: 10.1056/NEJMoa2103695; pmid: 34096690
15. D. S. Hong *et al*., $KRAS^{G12C}$ inhibition with sotorasib in advanced solid tumors. *N. Engl. J. Med.* **383**, 1207–1217 (2020). doi: 10.1056/NEJMoa1917239; pmid: 32955176
16. P. A. Jänne *et al*., Adagrasib in non-small-cell lung cancer harboring a $KRAS^{G12C}$ mutation. *N. Engl. J. Med.* **387**, 120–131 (2022). doi: 10.1056/NEJMoa2204619; pmid: 35658005
17. N. S. Akhave, A. B. Biter, D. S. Hong, Mechanisms of resistance to $KRAS^{G12C}$-targeted therapy. *Cancer Discov.* **11**, 1345–1352 (2021). doi: 10.1158/2159-8290.CD-20-1616; pmid: 33820777
18. N. Tanaka *et al*., Clinical acquired resistance to KRASG12C inhibition through a novel KRAS switch-II pocket mutation and polyclonal alterations converging on RAS-MAPK reactivation. *Cancer Discov.* **11**, 1913–1922 (2021). doi: 10.1158/2159-8290.CD-21-0365; pmid: 33824136
19. M. M. Awad *et al*., Acquired resistance to $KRAS^{G12C}$ inhibition in cancer. *N. Engl. J. Med.* **384**, 2382–2393 (2021). doi: 10.1056/NEJMoa2105281; pmid: 34161704
20. B. T. Li *et al*., Largest evaluation of acquired resistance to sotorasib in KRAS p.G12C-mutated non–small cell lung cancer (NSCLC) and colorectal cancer (CRC): Plasma biomarker analysis of CodeBreaK100. *J. Clin. Oncol.* **40** (16_suppl), 102–102 (2022). doi: 10.1200/JCO.2022.40.16_suppl.102
21. Y. Zhao *et al*., Diverse alterations associated with resistance to KRAS(G12C) inhibition. *Nature* **599**, 679–683 (2021). doi: 10.1038/s41586-021-04065-2; pmid: 34759319
22. R. Yaeger *et al*., Molecular characterization of acquired resistance to KRASG12C-EGFR inhibition in colorectal cancer. *Cancer Discov.* **13**, 41–55 (2023). doi: 10.1158/2159-8290.CD-22-0405; pmid: 36355783
23. Y. S. Tsai *et al*., Rapid idiosyncratic mechanisms of clinical resistance to KRAS G12C inhibition. *J. Clin. Invest.* **132**, e155523 (2022). doi: 10.1172/JCI155523; pmid: 34990404
24. V. Amodio *et al*., EGFR blockade reverts resistance to KRAS G12C inhibition in colorectal cancer. *Cancer Discov.* **10**, 1129–1139 (2020). doi: 10.1158/2159-8290.CD-20-0187; pmid: 32430388
25. J. Y. Xue *et al*., Rapid non-uniform adaptation to conformation-specific KRAS(G12C) inhibition. *Nature* **577**, 421–425 (2020). doi: 10.1038/s41586-019-1884-x; pmid: 31915379
26. M. B. Ryan *et al*., Vertical pathway inhibition overcomes adaptive feedback resistance to $KRAS^{G12C}$ inhibition. *Clin. Cancer Res.* **26**, 1633–1643 (2020). doi: 10.1158/1078-0432.CCR-19-3523; pmid: 31776128
27. W. S. Brown *et al*., Overcoming adaptive resistance to KRAS and MEK inhibitors by co-targeting mTORC1/2 complexes in pancreatic cancer. *Cell Rep. Med.* **1**, 100131 (2020). doi: 10.1016/j.xcrm.2020.100131; pmid: 33294856
28. M. Molina-Arcas *et al*., Development of combination therapies to maximize the impact of KRAS-G12C inhibitors in lung cancer. *Sci. Transl. Med.* **11**, eaaw7999 (2019). doi: 10.1126/scitranslmed.aaw7999; pmid: 31534020
29. S. Suzuki *et al*., KRAS inhibitor resistance in *MET*-amplified $KRAS^{G12C}$ non-small cell lung cancer induced by RAS- and non-RAS-mediated cell signaling mechanisms. *Clin. Cancer Res.* **27**, 5697–5707 (2021). doi: 10.1158/1078-0432.CCR-21-0856; pmid: 34365406
30. S. R. Punekar, V. Velcheti, B. G. Neel, K. K. Wong, The current state of the art and future trends in RAS-targeted cancer therapies. *Nat. Rev. Clin. Oncol.* **19**, 637–655 (2022). doi: 10.1038/s41571-022-00671-9; pmid: 36028717
31. D. L. Longo, N. Rosen, Targeting oncogenic RAS protein. *N. Engl. J. Med.* **387**, 184–186 (2022). doi: 10.1056/NEJMe2206831; pmid: 35830646
32. V. Thatikonda *et al*., Combined $KRAS^{G12C}$ and SOS1 inhibition enhances and extends the anti-tumor response in $KRAS^{G12C}$-driven cancers by addressing intrinsic and acquired resistance. biorXiv 525210 [Preprint] (2023); https://doi.org/10.1101/2023.01.23.525210.
33. C. Fedele *et al*., SHP2 inhibition diminishes KRASG12C cycling and promotes tumor microenvironment remodeling. *J. Exp. Med.* **218**, e20201414 (2021). doi: 10.1084/jem.20201414; pmid: 33045063
34. C. S. L. Ho *et al*., HER2 mediates clinical resistance to the $KRAS^{G12C}$ inhibitor sotorasib, which is overcome by co-targeting SHP2. *Eur. J. Cancer* **159**, 16–23 (2021). doi: 10.1016/j.ejca.2021.10.003; pmid: 34715459
35. M. B. Ryan *et al*., $KRAS^{G12C}$-independent feedback activation of wild-type RAS constrains $KRAS^{G12C}$ inhibitor efficacy. *Cell Rep.* **39**, 110993 (2022). doi: 10.1016/j.celrep.2022.110993; pmid: 35732135
36. M. Sattler, A. Mohanty, P. Kulkarni, R. Salgia, Precision oncology provides opportunities for targeting KRAS-inhibitor resistance. *Trends Cancer* **9**, 42–54 (2023). doi: 10.1016/j.trecan.2022.10.001; pmid: 36751115
37. A. Kapoor *et al*., Yap1 activation enables bypass of oncogenic Kras addiction in pancreatic cancer. *Cell* **158**, 185–197 (2014). doi: 10.1016/j.cell.2014.06.003; pmid: 24954535
38. D. D. Shao *et al*., KRAS and YAP1 converge to regulate EMT and tumor survival. *Cell* **158**, 171–184 (2014). doi: 10.1016/j.cell.2014.06.004; pmid: 24954536
39. G. H. Fisher *et al*., Induction and apoptotic regression of lung adenocarcinomas by regulation of a K-Ras transgene in the presence and absence of tumor suppressor genes. *Genes Dev.* **15**, 3249–3262 (2001). doi: 10.1101/gad.947701; pmid: 11751631
40. S. Joshi *et al*., Adapting to stress—Chaperome networks in cancer. *Nat. Rev. Cancer* **18**, 562–575 (2018). doi: 10.1038/s41568-018-0020-9; pmid: 29795326
41. X. Chen, J. R. Cubillos-Ruiz, Endoplasmic reticulum stress signals in the tumour and its microenvironment. *Nat. Rev. Cancer* **21**, 71–88 (2021). doi: 10.1038/s41568-020-00312-2; pmid: 33214692
42. J. Labbadia, R. I. Morimoto, The biology of proteostasis in aging and disease. *Annu. Rev. Biochem.* **84**, 435–464 (2015). doi: 10.1146/annurev-biochem-060614-033955; pmid: 25784053
43. P. Walter, D. Ron, The unfolded protein response: From stress pathway to homeostatic regulation. *Science* **334**, 1081–1086 (2011). doi: 10.1126/science.1209038; pmid: 22116877
44. M. Wang, R. J. Kaufman, Protein misfolding in the endoplasmic reticulum as a conduit to human disease. *Nature* **529**, 326–335 (2016). doi: 10.1038/nature17041; pmid: 26791723
45. A. Almanza *et al*., Endoplasmic reticulum stress signalling - from basic mechanisms to clinical applications. *FEBS J.* **286**, 241–278 (2019). doi: 10.1111/febs.14608; pmid: 30027602
46. R. Gomez-Pastor, E. T. Burchfiel, D. J. Thiele, Regulation of heat shock transcription factors and their roles in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **19**, 4–19 (2018). doi: 10.1038/nrm.2017.73; pmid: 28852220
47. Z. Tang *et al*., MEK guards proteome stability and inhibits tumor-suppressive amyloidogenesis via HSF1. *Cell* **160**, 729–744 (2015). doi: 10.1016/j.cell.2015.01.028; pmid: 25679764
48. H. J. Clarke, J. E. Chambers, E. Liniker, S. J. Marciniak, Endoplasmic reticulum stress in malignancy. *Cancer Cell* **25**, 563–573 (2014). doi: 10.1016/j.ccr.2014.03.015; pmid: 24823636
49. J. S. Cox, C. E. Shamu, P. Walter, Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. *Cell* **73**, 1197–1206 (1993). doi: 10.1016/0092-8674(93)90648-A; pmid: 8513503
50. K. Mori, W. Ma, M.-J. Gething, J. Sambrook, A transmembrane protein with a cdc2+/CDC28-related kinase activity is required for signaling from the ER to the nucleus. *Cell* **74**, 743–756 (1993). doi: 10.1016/0092-8674(93)90521-Q; pmid: 8358794
51. H. Yoshida, T. Matsui, A. Yamamoto, T. Okada, K. Mori, XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell* **107**, 881–891 (2001). doi: 10.1016/S0092-8674(01)00611-0; pmid: 11779464
52. M. Calfon *et al*., IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. *Nature* **415**, 92–96 (2002). doi: 10.1038/415092a; pmid: 11780124
53. C. Denoyelle *et al*., Anti-oncogenic role of the endoplasmic reticulum differentially activated by mutations in the MAPK pathway. *Nat. Cell Biol.* **8**, 1053–1063 (2006). doi: 10.1038/ncb1471; pmid: 16964246
54. T. Šuštić *et al*., A role for the unfolded protein response stress sensor ERN1 in regulating the response to MEK inhibitors in KRAS mutant colon cancers. *Genome Med.* **10**, 90 (2018). doi: 10.1186/s13073-018-0600-z; pmid: 30482246
55. H. Ying *et al*., Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* **149**, 656–670 (2012). doi: 10.1016/j.cell.2012.01.058; pmid: 22541435
56. M. R. Nilsson, Techniques to study amyloid fibril formation in vitro. *Methods* **34**, 151–160 (2004). doi: 10.1016/j.ymeth.2004.03.012; pmid: 15283924
57. M. Galves, R. Rathi, G. Prag, A. Ashkenazi, Ubiquitin signaling and degradation of aggregate-prone proteins. *Trends Biochem. Sci.* **44**, 872–884 (2019). doi: 10.1016/j.tibs.2019.04.007; pmid: 31079890
58. C. L. Klaips, G. G. Jayaraj, F. U. Hartl, Pathways of cellular proteostasis in aging and disease. *J. Cell Biol.* **217**, 51–63 (2018). doi: 10.1083/jcb.201709072; pmid: 29127110
59. L. Cortez, V. Sim, The therapeutic potential of chemical chaperones in protein folding diseases. *Prion* **8**, 197–202 (2014). doi: 10.4161/pri.28938; pmid: 24818993
60. F. Engin *et al*., Restoration of the unfolded protein response in pancreatic β cells protects mice against type 1 diabetes. *Sci. Transl. Med.* **5**, 211ra156 (2013). doi: 10.1126/scitranslmed.3006534; pmid: 24225943
61. M. Costa-Mattioli, P. Walter, The integrated stress response: From mechanism to disease. *Science* **368**, eaat5314 (2020). doi: 10.1126/science.aat5314; pmid: 32327570
62. T. Guettouche, F. Boellmann, W. S. Lane, R. Voellmy, Analysis of phosphorylation of human heat shock factor 1 in cells experiencing a stress. *BMC Biochem.* **6**, 4 (2005). doi: 10.1186/1471-2091-6-4; pmid: 15760475
63. W. Tirasophon, A. A. Welihinda, R. J. Kaufman, A stress response pathway from the endoplasmic reticulum to the nucleus requires a novel bifunctional protein kinase/endoribonuclease (Ire1p) in mammalian cells. *Genes Dev.* **12**, 1812–1824 (1998). doi: 10.1101/gad.12.12.1812; pmid: 9637683
64. J. Hollien, J. S. Weissman, Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. *Science* **313**, 104–107 (2006). doi: 10.1126/science.1129631; pmid: 16825573
65. S. Lhomond *et al*., Dual IRE1 RNase functions dictate glioblastoma development. *EMBO Mol. Med.* **10**, e7929 (2018). doi: 10.15252/emmm.201707929; pmid: 29311133
66. C. Fedele *et al*., SHP2 inhibition prevents adaptive resistance to MEK inhibitors in multiple cancer models. *Cancer Discov.* **8**, 1237–1249 (2018). doi: 10.1158/2159-8290.CD-18-0444; pmid: 30045908
67. S. Mainardi *et al*., SHP2 is required for growth of KRAS-mutant non-small-cell lung cancer in vivo. *Nat. Med.* **24**, 961–967 (2018). doi: 10.1038/s41591-018-0023-9; pmid: 29808006
68. R. J. Nichols *et al*., RAS nucleotide cycling underlies the SHP2 phosphatase dependence of mutant BRAF-, NF1- and RAS-driven cancers. *Nat. Cell Biol.* **20**, 1064–1073 (2018). doi: 10.1038/s41556-018-0169-1; pmid: 30104724
69. D. A. Ruess *et al*., Mutant KRAS-driven cancers depend on PTPN11/SHP2 phosphatase. *Nat. Med.* **24**, 954–960 (2018). doi: 10.1038/s41591-018-0024-8; pmid: 29808009
70. S. Sun *et al*., IRE1α is an endogenous substrate of endoplasmic-reticulum-associated degradation. *Nat. Cell Biol.* **17**, 1546–1555 (2015). doi: 10.1038/ncb3266; pmid: 26551274
71. J. L. Brodsky, Cleaning up: ER-associated degradation to the rescue. *Cell* **151**, 1163–1167 (2012). doi: 10.1016/j.cell.2012.11.012; pmid: 23217703
72. X. Liu *et al*., Notch-induced endoplasmic reticulum-associated degradation governs mouse thymocyte β-selection. *eLife* **10**, e69975 (2021). doi: 10.7554/eLife.69975; pmid: 34240701
73. L. Xu *et al*., Protein quality control through endoplasmic reticulum-associated degradation maintains haematopoietic stem cell identity and niche interactions. *Nat. Cell Biol.* **22**, 1162–1169 (2020). doi: 10.1038/s41556-020-00581-x; pmid: 32958856
74. H. Meyer, M. Bug, S. Bremer, Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. *Nat. Cell Biol.* **14**, 117–123 (2012). doi: 10.1038/ncb2407; pmid: 22298039
75. D. J. Anderson *et al*., Targeting the AAA ATPase p97 as an approach to treat cancer through disruption of protein homeostasis. *Cancer Cell* **28**, 653–665 (2015). doi: 10.1016/j.ccell.2015.10.002; pmid: 26555175
76. R. Roskoski Jr., ERK1/2 MAP kinases: Structure, function, and regulation. *Pharmacol. Res.* **66**, 105–143 (2012). doi: 10.1016/j.phrs.2012.04.005; pmid: 22569528
77. S. Satpathy *et al*., A proteogenomic portrait of lung squamous cell carcinoma. *Cell* **184**, 4348–4371.e40 (2021). doi: 10.1016/j.cell.2021.07.016; pmid: 34358469
78. K. You *et al*., QRICH1 dictates the outcome of ER stress through transcriptional control of proteostasis. *Science* **371**, eabb6896 (2021). doi: 10.1126/science.abb6896; pmid: 33384352
79. C. Y. Liu, Z. Xu, R. J. Kaufman, Structure and intermolecular interactions of the luminal dimerization domain of human IRE1alpha. *J. Biol. Chem.* **278**, 17680–17687 (2003). doi: 10.1074/jbc.M300418200; pmid: 12637535
80. L. Buscail, B. Bournet, P. Cordelier, Role of oncogenic KRAS in the diagnosis, prognosis and treatment of pancreatic cancer. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 153–168 (2020). doi: 10.1038/s41575-019-0245-4; pmid: 32005945

81. A. W. Tolcher *et al*., Antitumor activity in RAS-driven tumors by blocking AKT and MEK. *Clin. Cancer Res.* **21**, 739–748 (2015). doi: 10.1158/1078-0432.CCR-14-1901; pmid: 25516890
82. V. Chung *et al*., Effect of selumetinib and MK-2206 vs oxaliplatin and fluorouracil in patients with metastatic pancreatic cancer after prior therapy: SWOG S1115 study randomized clinical trial. *JAMA Oncol.* **3**, 516–522 (2017). doi: 10.1001/jamaoncol.2016.5383; pmid: 27978579
83. N. Zhao *et al*., Pharmacological targeting of MYC-regulated IRE1/XBP1 pathway suppresses MYC-driven breast cancer. *J. Clin. Invest.* **128**, 1283–1299 (2018). doi: 10.1172/JCI95873; pmid: 29480818
84. P. J. Le Reste *et al*., Local intracerebral inhibition of IRE1 by MKC8866 sensitizes glioblastoma to irradiation/chemotherapy in vivo. *Cancer Lett.* **494**, 73–83 (2020). doi: 10.1016/j.canlet.2020.08.028; pmid: 32882336
85. S. E. Logue *et al*., Inhibition of IRE1 RNase activity modulates the tumor cell secretome and enhances response to chemotherapy. *Nat. Commun.* **9**, 3267 (2018). doi: 10.1038/s41467-018-05763-8; pmid: 30111846
86. N. Y. Gabrail *et al*., A phase 1/2 trial of ORIN1001, a first-in-class IRE1 inhibitor, in patients with advanced solid tumors. *J. Clin. Oncol.* **39** (15_suppl), 3080–3080 (2021). doi: 10.1200/JCO.2021.39.15_suppl.3080
87. M. Salmón *et al*., Kras oncogene ablation prevents resistance in advanced lung adenocarcinomas. *J. Clin. Invest.* **133**, e164413 (2023). doi: 10.1172/JCI164413; pmid: 36928090
88. H. C. Mahler, W. Friess, U. Grauschopf, S. Kiese, Protein aggregation: Pathways, induction factors and analysis. *J. Pharm. Sci.* **98**, 2909–2934 (2009). doi: 10.1002/jps.21566; pmid: 18823031
89. M. F. Rimawi *et al*., Early efficacy evaluation of ORIN1001, a first in class IRE1 alpha inhibitor, in advanced solid tumors. *J. Clin. Oncol.* **41** (16_suppl), 1092–1092 (2023). doi: 10.1200/JCO.2023.41.16_suppl.1092
90. M. P. Kim *et al*., Generation of orthotopic and heterotopic human pancreatic cancer xenografts in immunodeficient mice. *Nat. Protoc.* **4**, 1670–1680 (2009). doi: 10.1038/nprot.2009.171; pmid: 19876027
91. L. Yan *et al*., Targeting glucose metabolism sensitizes pancreatic cancer to MEK inhibition. *Cancer Res.* **81**, 4054–4065 (2021). doi: 10.1158/0008-5472.CAN-20-3792; pmid: 34117030
92. Z. Cheng *et al*., Inhibition of BET bromodomain targets genetically diverse glioblastoma. *Clin. Cancer Res.* **19**, 1748–1759 (2013). doi: 10.1158/1078-0432.CCR-12-3066; pmid: 23403638
93. N. E. Sanjana, O. Shalem, F. Zhang, Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014). doi: 10.1038/nmeth.3047; pmid: 25075903
94. P. Bankhead *et al*., QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017). doi: 10.1038/s41598-017-17204-5; pmid: 29203879
95. K. H. Wrighton *et al*., Small C-terminal domain phosphatases dephosphorylate the regulatory linker regions of Smad2 and Smad3 to enhance transforming growth factor-β signaling. *J. Biol. Chem.* **281**, 38365–38375 (2006). doi: 10.1074/jbc.M607246200; pmid: 17035229
96. B. C. Cross *et al*., The molecular basis for selective inhibition of unconventional mRNA splicing by an IRE1-binding small molecule. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E869–E878 (2012). doi: 10.1073/pnas.1115623109; pmid: 22315414
97. S. V. Vasaikar, P. Straub, J. Wang, B. Zhang, LinkedOmics: Analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res.* **46** (D1), D956–D963 (2018). doi: 10.1093/nar/gkx1090; pmid: 29136207

ACKNOWLEDGMENTS

We thank J. Rosen, Q. Zhang, and J. Xu for advice, discussion, and critical review of the manuscript. **Funding:** This work was supported by US Department of Defense Congressionally Directed Medical Research Programs (W81XWH1910524 to X.C., W81XWH1910035 to X. Lv), the National Institutes of Health (R01CA270240, R37CA228304, P50CA186784, and R01HL146642 to X.C.; R01CA214793 and 5P01CA117969 to H.Y.; 1R01CA272744 to W.Y.; R01DK115454 and R01GM142143 to A.C.; T32DK60445-17 to L.M.; CA215591 and SPORE DRP CA126752 to Z.S.; U24CA210954 to B.Z.; U54CA224065 and U54CA224081 to J.A.R.; R01GM130838 and R01NS117668 to Y.C.; and R01HD007857 and R01HD008188 to B.W.O.), Cancer Prevention and Research Institute of Texas (RP230285 to X.C., RR140038 CPRIT Scholar in Cancer Research award to A.C., RP160283 Baylor College of Medicine Comprehensive Cancer Training Program award to F.P., and RR160027 CPRIT Scholar in Cancer Research award to B.Z.), and American Cancer Society (RSG-22-017-01-CCB to W.Y.). This work used the Aperio GT-450 digital pathology scanner in the Baylor College of Medicine Breast Center Pathology Core that was purchased with funding from the NIH S10 grant S10OD028671. This work was supported by the Cytometry and Cell Sorting Core at the Baylor College of Medicine with funding from the CPRIT Core Facility Support Award (CPRIT-RP180672), the NIH (CA125123, S10OD025251, and RR024574), and the assistance of J. M. Sederstrom. Imaging for this work was supported by the Integrated Microscopy Core at the Baylor College of Medicine with funding from NIH (DK56338 and CA125123) and CPRIT (RP150578). This work was supported by the DNA Sequencing Core at the Baylor College of Medicine with funding from the NIH support (R01HD100535) and the assistance of Y. Lei and P. Pimienta Ramirez. **Author contributions:** X.C., H.Y., X. Lv, and X. Lu conceived the project and designed the research. X. Lv, X. Lu, J.C., Y.D., Q.L., Fa.P., D.L., D.H., E.M., D.W.C., X.W., S.M., J.Y., Fe.P., L.Y., L.M., and Y.H. performed the experiments and analyzed the data. H.M., A.P., Y.C., W.Y., E.C.C., A.C., X.Li., G.M., B.B., J.W., P.H., Z.S., H.W., J.A.R., B.W.O., Q.C.Y., M.J.E., and M.F.R. contributed to discussions, experimental design, and critical reagents. S.R.S. and B.Z. analyzed the CPTAC data. X.C. supervised the project. X.C., X. Lv, and X. Lu wrote the paper. **Competing interests:** M.F.R. receives research funding from Pfizer and Genentech. M.F.R. is a consultant and receives consulting fees from: Novartis, Seagen, Macrogenics, and AstraZeneca. M.F.R. is the principal investigator of the clinical trial NCT03950570. X.C. reports previous research funding from Fosun Pharma. J.A.R. reports receiving a commercial research grant from The University of Texas MD Anderson Cancer Center, has a sponsored research agreement from Genprex, has ownership interest (including stock, patents, etc.) in Genprex, and is a consultant and advisory board member for Genprex. M.J.E. reports full-time employment with AstraZeneca from 7 March 2022. M.J.E. holds patents and receives income from Veracyte on the PAM50-based products, Prosigna. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. Requests for materials should be addressed to X.C.
License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn4180
Figs. S1 to S18
Tables S1 to S5
MDAR Reproducibility Checklist

Submitted 24 November 2021; resubmitted 6 March 2023
Accepted 28 July 2023
10.1126/science.abn4180

RESEARCH ARTICLE SUMMARY

IMMUNOLOGY

Kupffer cell–like syncytia replenish resident macrophage function in the fibrotic liver

Moritz Peiseler†, Bruna Araujo David†, Joel Zindel, Bas G. J. Surewaard, Woo-Yong Lee, Felix Heymann, Ysbrand Nusse, Fernanda V. S. Castanheira, Raymond Shim, Adrien Guillot, Alix Bruneau, Jawairia Atif, Catia Perciani, Christina Ohland, Priyanka Ganguli Mukherjee, Annika Niehrs, Roland Thuenauer, Marcus Altfeld, Mathias Amrein, Zhaoyuan Liu, Paul M. K. Gordon, Kathy McCoy, Justin Deniset, Sonya MacParland, Florent Ginhoux, Frank Tacke, Paul Kubes*

INTRODUCTION: The local environment is critical for establishing the phenotype of macrophages within a given organ. In the liver, resident macrophages reach out of the sinusoids to receive instructive cues in a niche composed of hepatocytes, endothelial cells, and stellate cells. These cues activate specific transcription factors, which endow these macrophages with Kupffer cell (KC) "identity." In the sinusoids, KCs perform the critical function of capturing pathogens from the blood by means of specialized receptors, including the complement receptor CRIg. Liver fibrosis and cirrhosis is the common end-stage of various chronic liver diseases, leading to substantial morbidity and mortality in affected individuals. Despite different etiologies, progression is similar, with hepatocyte death and collagen deposition around sinusoids, resulting in the redistribution of blood flow to new and expanded intrahepatic and extrahepatic collateral vessels.

RATIONALE: It remains unclear how fibrotic remodeling of the niche environment affects the KC compartment. In this study, we used various lineage-tracing models and intravital microscopy to visualize, track, and functionally assess monocytes and KCs in the fibrotic liver environment.

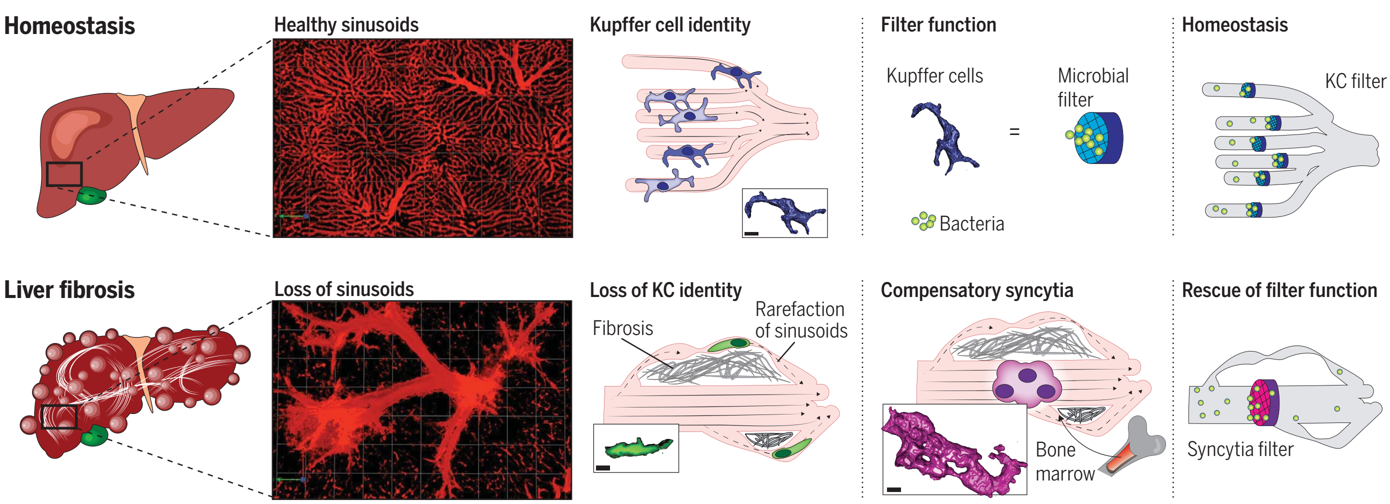
RESULTS: Using the most common mouse model of liver fibrosis—carbon tetrachloride toxicity—we observed profound architectural changes in the liver. This remodeling included a massive increase in collateral vessels and collagen deposition around the sinusoids, which caused KCs to lose contact with their surrounding environment. This, in turn, led to the down-regulation of key transcription factors and membrane proteins such as CLEC4F, CRIg, and TIM-4, which collectively determine KC identity. Although these changes resulted in impaired KC function, the liver continued to act as a major filter of blood-borne bacteria despite the loss of KC identity.

An abundance of monocytes were recruited, and these cells primarily adhered to large intrahepatic vessels through CD44 owing to increased endothelial cell adhesivity driven by intestinal microbiota. Monocytes form large clusters within the collateral vessels and began to express KC markers. These monocytes made up a spectrum of structures ranging from clusters of individual cells to fused multinucleated giant cells that collectively appeared as KC-like syncytia. Although individual KCs could not catch bacteria flowing within larger vessels, KC-like syncytia were able to capture high numbers of circulating bacteria. Using transcriptomic analysis, we identified CD36 as the key molecule underlying syncytial fusion and reduced susceptibility to infection. CRIg-expresssing intravascular macrophage syncytia were also found in human cirrhosis of different etiology.

CONCLUSION: Loss of contact with parenchymal cells in the fibrotic niche leads sinusoid-resident KCs to lose identity and function. Because KC replenishment in rarefied sinusoids would serve little purpose, monocytes follow the formation of collateral vessels that bypass sinusoids, where they form KC-like syncytia that have the capacity to capture bacteria from the bloodstream. Thus, KC maladaptation within an altered fibrotic niche environment is rescued by monocytes forming KC-like syncytia to capture bacteria. These cell structures may play a critical evolutionary role that allows mammals to withstand severe chronic insults in the liver. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: pkubes@ucalgary.ca
†These authors contributed equally to this work.
Cite this article as M. Peiseler *et al.*, *Science* **381**, eabq5202 (2023). DOI: 10.1126/science.abq5202

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abq5202



Kupffer cell adaptation in fibrotic liver disease. In healthy livers, KCs reside in sinusoids and rapidly capture blood-borne pathogens. In liver fibrosis, sinusoids are rarefied, leading to redistributed circulation through high-flow collateral vessels. KCs consequently lose their identity and function. Monocytes then seed larger vessels and form KC-like syncytia with the heightened capacity to capture bacteria, providing a rescue adaptation to the fibrosis-induced loss of KCs.

RESEARCH ARTICLE

IMMUNOLOGY

Kupffer cell–like syncytia replenish resident macrophage function in the fibrotic liver

Moritz Peiseler[1,2,3,4]†, Bruna Araujo David[1,2]†, Joel Zindel[1,2,5], Bas G. J. Surewaard[1,2], Woo-Yong Lee[1,2], Felix Heymann[3], Ysbrand Nusse[1,2], Fernanda V. S. Castanheira[1,2], Raymond Shim[1,2], Adrien Guillot[3], Alix Bruneau[3], Jawairia Atif[6], Catia Perciani[6], Christina Ohland[1,2], Priyanka Ganguli Mukherjee[7], Annika Niehrs[8], Roland Thuenauer[8], Marcus Altfeld[8], Mathias Amrein[7], Zhaoyuan Liu[9], Paul M. K. Gordon[10], Kathy McCoy[1,2,11], Justin Deniset[1,2,12,13], Sonya MacParland[6], Florent Ginhoux[14,15], Frank Tacke[3], Paul Kubes[1,2]*

Kupffer cells (KCs) are localized in liver sinusoids but extend pseudopods to parenchymal cells to maintain their identity and serve as the body's central bacterial filter. Liver cirrhosis drastically alters vascular architecture, but how KCs adapt is unclear. We used a mouse model of liver fibrosis and human tissue to examine immune adaptation. Fibrosis forced KCs to lose contact with parenchymal cells, down-regulating "KC identity," which rendered them incapable of clearing bacteria. Commensals stimulated the recruitment of monocytes through CD44 to a spatially distinct vascular compartment. There, recruited monocytes formed large aggregates of multinucleated cells (syncytia) that expressed phenotypical KC markers and displayed enhanced bacterial capture ability. Syncytia formed via CD36 and were observed in human cirrhosis as a possible antimicrobial defense that evolved with fibrosis.

Positioned at the confluence of the intestinal vasculature and systemic circulation, the liver acts as a major filter of disseminated blood-borne pathogens (*1*). Liver sinusoids are microanatomical filtration units designed for the optimal surveillance of passing blood (*2*). Strategically located within these sinusoids is a population of resident macrophages called Kupffer cells (KCs), which represent the most abundant population of tissue macrophages in the human body (*3*). KCs are characterized by their ramified structures, sessile nature relative to other patrolling immune cells, and primarily intravascular location (Fig. 1A, fig. S1A, and movie S1). At the same time, their long protrusions interact with liver parenchymal cells (Fig. 1, A and B) including hepatic stellate cells (HSCs), liver sinusoidal endothelial cells (LSECs), and hepatocytes, together forming the KC niche. The interaction between parenchymal cells and KCs is critical for KC function, as parenchymal cells supply important signals that maintain KC identity (*4*). This niche is optimized for immune surveillance, as blood-borne pathogens are instantaneously captured by KCs (Fig. 1C and movie S2).

[1]Department of Physiology and Pharmacology, University of Calgary, Calgary, Alberta, Canada. [2]Snyder Institute for Chronic Diseases, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada. [3]Department of Hepatology and Gastroenterology, Charité Universitätsmedizin Berlin, Campus Virchow Klinikum and Campus Charité Mitte, Berlin, Germany. [4]Berlin Institute of Health (BIH), Berlin, Germany. [5]Department of Visceral Surgery and Medicine, Department for BioMedical Research (DBMR), University of Bern, Bern, Switzerland. [6]Ajmera Transplant Centre, Toronto General Research Institute, University Health Network, Toronto, Ontario, Canada. [7]Department of Cell Biology and Anatomy, University of Calgary, Calgary, Alberta, Canada. [8]Leibniz Institute of Virology (LIV), Hamburg, Germany. [9]Shanghai Institute of Immunology, Department of Immunology and Microbiology, Shanghai Jiao Tong University School of Medicine, Shanghai, China. [10]Centre for Health Genomics and Informatics, University of Calgary, Calgary, Alberta, Canada. [11]Department of Microbiology, Immunology and Infectious Diseases, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada. [12]Department of Cardiac Sciences, University of Calgary, Calgary, Alberta, Canada. [13]Libin Cardiovascular Institute, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada. [14]Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore. [15]Gustave Roussy Cancer Campus, INSERM U1015, Villejuif, France.
***Corresponding author. Email: pkubes@ucalgary.ca**
†These authors contributed equally to this work.

Intravital imaging reveals progressive vascular remodeling in liver fibrosis

In order for KCs to capture blood-borne bacteria, sinusoids must be sufficiently narrow so that KCs occupy large amounts of the sinusoidal lumen without restricting blood flow through these channels (Fig. 1C). In chronic liver diseases with fibrosis, architectural changes—regardless of underlying etiology—occur, including the deposition of a basement membrane around sinusoids (capillarization), defenestration of sinusoids, and the development of intrahepatic and extrahepatic shunts and collaterals to accommodate blood flow (*5*–*7*). Using intravital microscopy (IVM), we imaged numerous mouse models of liver disease. We found that the carbon tetrachloride (CCl_4) toxicity model produced the most pronounced fibrosis and vascular changes like those seen in human disease, allowing us to study KC immunodynamics in a bona fide fibrotic niche (Fig. 2A). IVM revealed profound parenchymal changes that included narrowing of liver sinusoids and dilation of larger venules (Fig. 2A and fig. S1, B and C). Imaging of blood flow revealed the homogeneous perfusion of sinusoids in control livers (fig. S1D and movie S3), whereas livers of CCl_4-treated mice at 8 weeks showed large collateral vessels and dilated venules with poor or absent perfusion of sinusoids (fig. S1D and movie S3), which are characteristic of human cirrhosis (*8*–*10*).

Tracking labeled red blood cells (RBCs) allowed us to quantify hepatic blood flow (Fig. 2, B to D). In control mice, homogeneous sinusoidal blood had a mean flow rate of 52 μm/s, whereas flow in the constricted sinusoids of fibrotic mice was lower (mean flow: 17 μm/s) and often completely occluded (Fig. 2, B and C). In venules, mean flow in control mice was ~200 μm/s, whereas mean flow through the collateral vessels of fibrotic mice was 350 μm/s (Fig. 2, B and D). The flow was often turbulent and not laminar (i.e., RBCs no longer traveled in straight lines) (Fig. 2D). The drastic increase of collateral vessels due to loss of patent capillaries is consistent with a threefold increase in venular cells (*11*). Furthermore, single-cell sequencing previously revealed that *ACKR1*, a venular endothelium signature gene, is highly enriched in human cirrhosis (*12*).

Fibrotic remodeling of the Kupffer cell niche

Second-harmonic generation revealed robust collagen deposition around the sinusoids of CCl_4-treated mice but not sham-treated mice (Fig. 2A), which was not easily seen by conventional histology (fig. S1, E and F). KCs in healthy mice were found exclusively in hepatic sinusoids and projected numerous pseudopods into the space of Disse (Fig. 2, E and F). By contrast, after 8 weeks, CCl_4-treated mice showed increased fibrosis and sinusoids of reduced diameter harboring macrophages with elongated shapes and significantly fewer pseudopods (Fig. 2, E and F). Patients with liver cirrhosis (stage F4 fibrosis) similarly presented with KCs that had lower average cell size and elongated shape compared with healthy individuals (Fig. 2, G and H, and fig. S1G). Furthermore, IVM of *Lrat*^{cre}×*Rosa26*^{tdTomato} HSC reporter mice (*13*) uncovered close contacts between KCs and HSCs in control mice (Fig. 1B and fig. S1H). Colocalization analysis confirmed the intricate anatomical interconnections between KCs and HSCs under homeostatic conditions (Fig. 2I and fig. S1I). By contrast, most KCs in the sinusoids of CCl_4-treated mice were detached from HSCs (Fig. 2I and fig. S1H), as HSCs relocated away from sinusoids and surrounded the large collaterals (fig. S1H), further altering the KC niche.

Macrophages in fibrotic sinusoids are functionally impaired

To investigate the bacterial capture of KCs in sinusoids, we induced bacteremia by intravenous

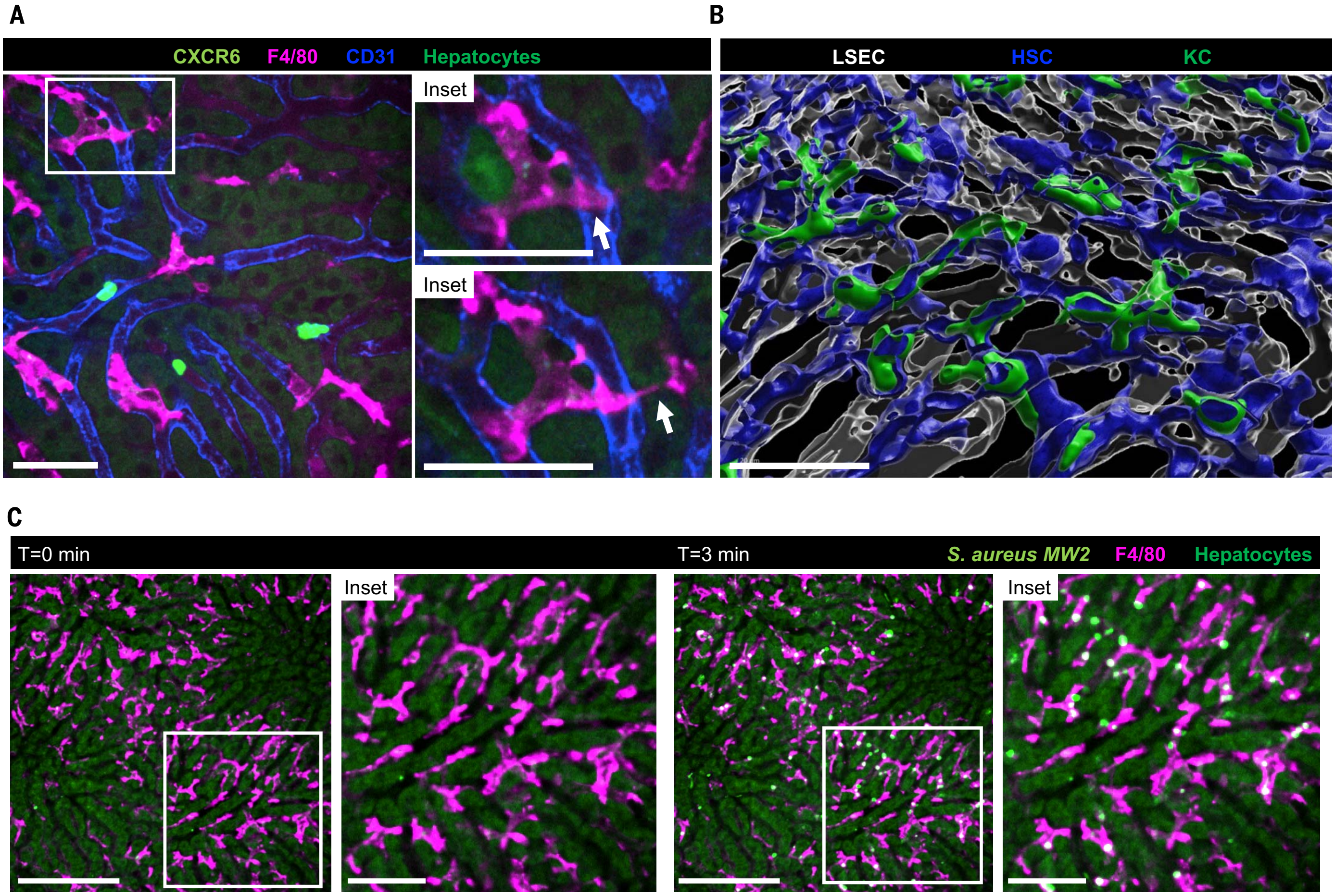


Fig. 1. Immune surveillance by Kupffer cells. (**A**) Representative IVM image of KCs (magenta) scanning the blood stream with their protrusions. Hepatocytes (dull green), LSECs (blue), and patrolling CXCR6-expressing cells (bright green) are depicted. (Inset) Time-lapse image of a KC sending a protrusion laterally touching another KC (arrow). Scale bars: 50 μm. (**B**) Representative 3D surface rendering of the KC niche showing KCs (green), hepatic stellate cells (blue), and endothelial cells (white). Scale bar: 100 μm. (**C**) Representative IVM image of staphylococcal capture (MW2-GFP in bright green) by KCs labeled with F4/80 (magenta) and hepatocytes (dull green) in baseline mice before (left) and 3 min after injection of bacteria showing instant capture of bacteria. Scale bars: 150 μm and 50 μm (insets).

injection of a clinically important methicillin-resistant *Staphylococcus aureus* (MRSA) strain expressing a fluorescent reporter (*S. aureus* MW2). In control mice, the vast majority of injected *S. aureus* MW2 were rapidly captured by KCs in the liver sinusoids (Fig. 2J, fig. S1J, and movie S4). By contrast, bacterial capture in the sinusoids of fibrotic mice was reduced (Fig. 2J, fig. S1J, and movie S4). The bacterial burden in the blood and spleens of CCl_4-treated mice was also higher than in controls (Fig. 2K). Moreover, the frequency of F4/80$^+$ macrophages in sinusoids that had captured bacteria decreased in fibrosis (Fig. 2L).

Given that KCs also kill the captured staphylococci (*14*), we next performed intravital imaging of the phagocytic processing capacity of sinusoidal KCs in control and CCl_4-treated mice (*15*). In control mice, the vast majority of captured bioparticles were rapidly acidified (>95%) (Fig. 2M and movie S5), whereas in CCl_4-treated mice, bioparticle acidification was delayed over the first hour (Fig. 2M and movie S5), suggesting the loss of KC antimicrobial capacity.

Macrophages in fibrotic sinusoids lose Kupffer cell identity

The morphological and functional changes in the KC compartment in fibrotic sinusoids prompted us to investigate phenotypical adaptation. We focused on two key molecules—complement receptor of the immunoglobulin superfamily [CRIg, also referred to as V-set immunoglobulin-domain-containing 4 (VSIG4)] and T cell membrane protein 4 (TIM-4)—that give KCs their functional phenotype (*16*). More than 90% of the F4/80$^+$ macrophages in the liver sinusoids of control mice coexpressed TIM-4 and CRIg (Fig. 3, A and B). By contrast, fewer F4/80$^+$ macrophages from CCl_4-treated mice expressed TIM-4 and/or CRIg (Fig. 3, A and B). These cells were all located within the sinusoids and, by definition, were KCs. F4/80$^+$ cells expressing TIM-4 and/or CRIg showed similar cell volumes compared with KCs in control mice, whereas F4/80$^+$ macrophages that did not express either marker were smaller (fig. S2A). All macrophage subsets in fibrotic sinusoids showed similar reductions in contact with HSCs compared with KCs from control sinusoids (fig. S2B).

Using multiparametric flow cytometry and dimensionality reduction with t-distributed stochastic neighbor embedding (tSNE) to cluster hepatic macrophages (fig. S2C), we found that most KCs in control mice expressed high levels of the known KC identity markers C-type lectin like receptor 2 (CLEC2), C-type lectin domain family 4 member F (CLEC4F), CRIg, and TIM-4 (Fig. 3C). By contrast, KCs in mice treated with CCl_4 for 8 weeks consistently had lower levels of all four molecules (Fig. 3C).

CRIg identifies a macrophage subset in fibrotic liver sinusoids with preserved bacterial capture phenotype

CRIg (VSIG4) is a complement receptor that is critical for KC-mediated bacterial capture (*17*, *18*). Indeed, in both control and CCl_4-treated mice, F4/80$^+$CRIg$^+$ and F4/80$^+$CRIg$^+$TIM-4$^+$

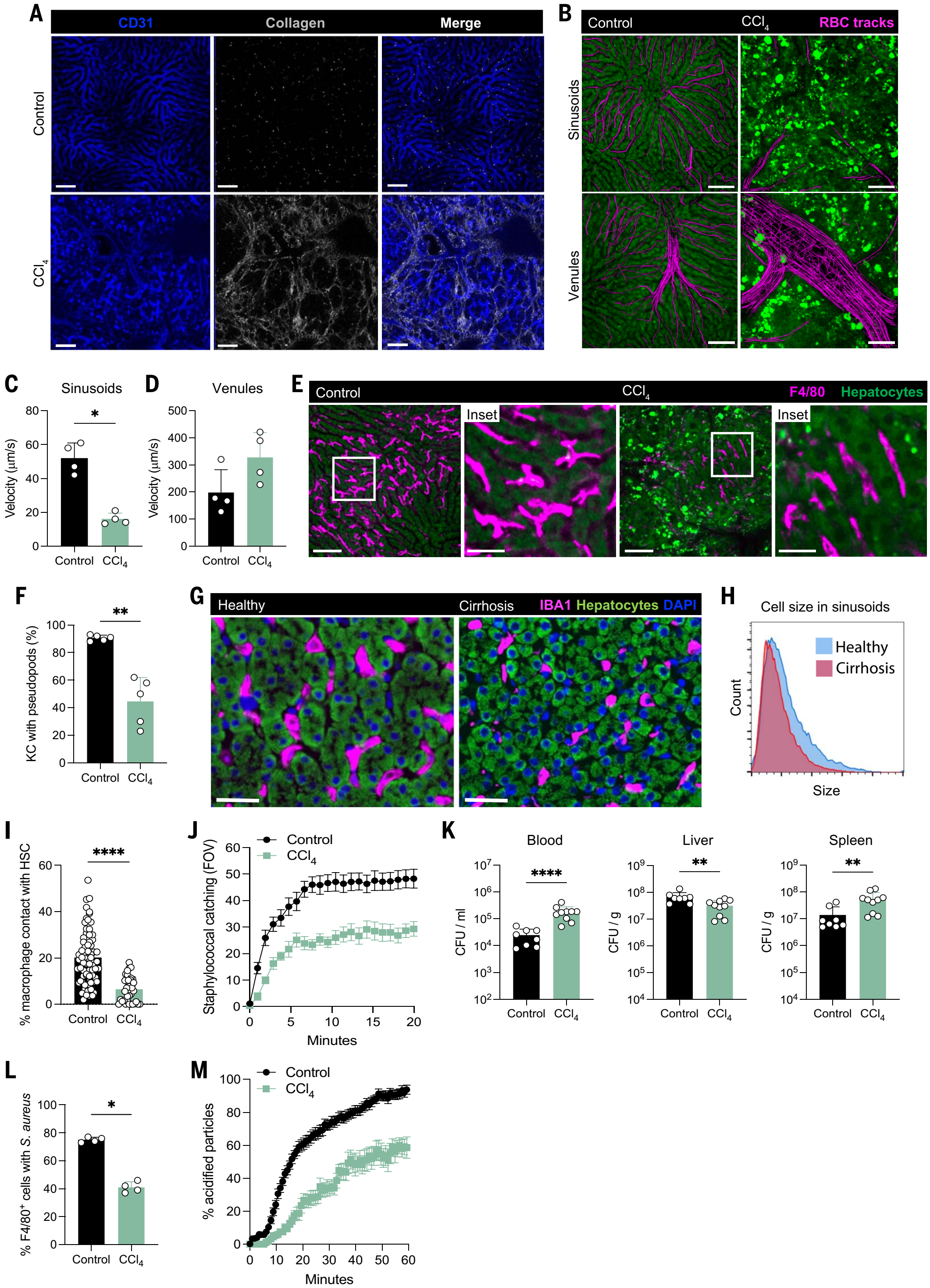


Fig. 2. Liver fibrosis alters the Kupffer cell niche resulting in a loss of function. (**A**) Representative multiphoton IVM stitched images of control and fibrotic liver, depicting CD31 (blue) and second-harmonic generation (SHG, white). Scale bars: 200 μm. (**B**) Representative IVM images of sinusoids and venules in control and CCl_4-treated mice. RBC tracks visualized (magenta) next to hepatocytes (dull green). Scale bars: 75 μm. (**C** and **D**) Mean velocity of RBC in sinusoids and venules (n = 4 per group, N = 3). (**E**) Representative IVM image of KC morphology. KCs (magenta), hepatocytes (dull green), and liver autofluorescence (bright green) Scale bars: 100 μm and 25 μm (insets). (**F**) Quantification of KCs with pseudopods in sinusoids (n = 5 per group, N = 3). (**G**) Representative morphology of IBA1$^+$ cells (magenta) in human liver cirrhosis. Nuclei (blue) and hepatocytes (green) are depicted. Scale bars: 50 μm. (**H**) Individual cell area of IBA1$^+$ cells located in the sinusoids of a healthy (blue) and cirrhotic liver (red). (**I**) Percentage of contact between *Lrat*-expressing HSCs and F4/80$^+$ macrophages in sinusoids (n = 3 per group, N = 3). (**J**) Quantification of bacterial capture in control and fibrotic mice infected with 5 × 10^7 CFU *S. aureus* MW2 from a 20-min video (n = 3 or 4 per group, N = 3 or 4). (**K**) Control and CCl_4-treated mice were infected with 5 × 10^7 CFU *S. aureus* MW2, and bacterial loads were determined 30 min after infection (n = 8 to 10 per group, N = 3 or 4). (**L**) Frequency of F4/80$^+$ cells in sinusoids with bacteria. (**M**) Quantification of acidification of pHrodo *S. aureus* bioparticles over 60-min video (n = 6 per group, N = 2). Data represented as individual values with mean ± SD and as mean ± SEM in (J) and (M). Mann-Whitney test used for (C), (D), (F), (I), (K), and (L). $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$.

macrophages captured bacteria, whereas only a small fraction of F4/80$^+$CRIg$^-$ cells did so (Fig. 3D and movie S6). Furthermore, F4/80$^+$CRIg$^+$ TIM-4$^+$ and F4/80$^+$CRIg$^+$ TIM-4$^-$ cells exhibited a capture efficiency of around 70% (fig. S2D), which was similar to KCs from control mice (Fig. 2L). By contrast, F4/80$^+$CRIg$^-$ cells demonstrated very low capture efficiency (fig. S2D).

Ontogeny of hepatic macrophages in liver sinusoids

Recent single-cell studies in models of fatty liver disease suggest that KCs are replaced by recruited monocyte-derived cells on the basis of transcriptomic changes (*19*–*21*). To study the ontogeny of sinusoidal KCs, we used *Ms4a3*^{cre}×*Rosa*^{tdtomato} monocyte fate-mapping mice (*22*), whose bone marrow (BM)–derived precursors, but not resident fetal liver–derived macrophages, fluoresce. Under homeostatic conditions, 2 to 10% of BM-derived monocytes contributed to the pool of liver macrophages (Fig. 3, E and F, and fig. S2E). After 8 weeks of CCl_4 treatment, ~30% of macrophages in sinusoids were derived from BM-derived monocytes and 70% were fetal liver–derived KCs (Fig. 3, E and F, and fig. S2E). The labeling efficiency of Ms4a3

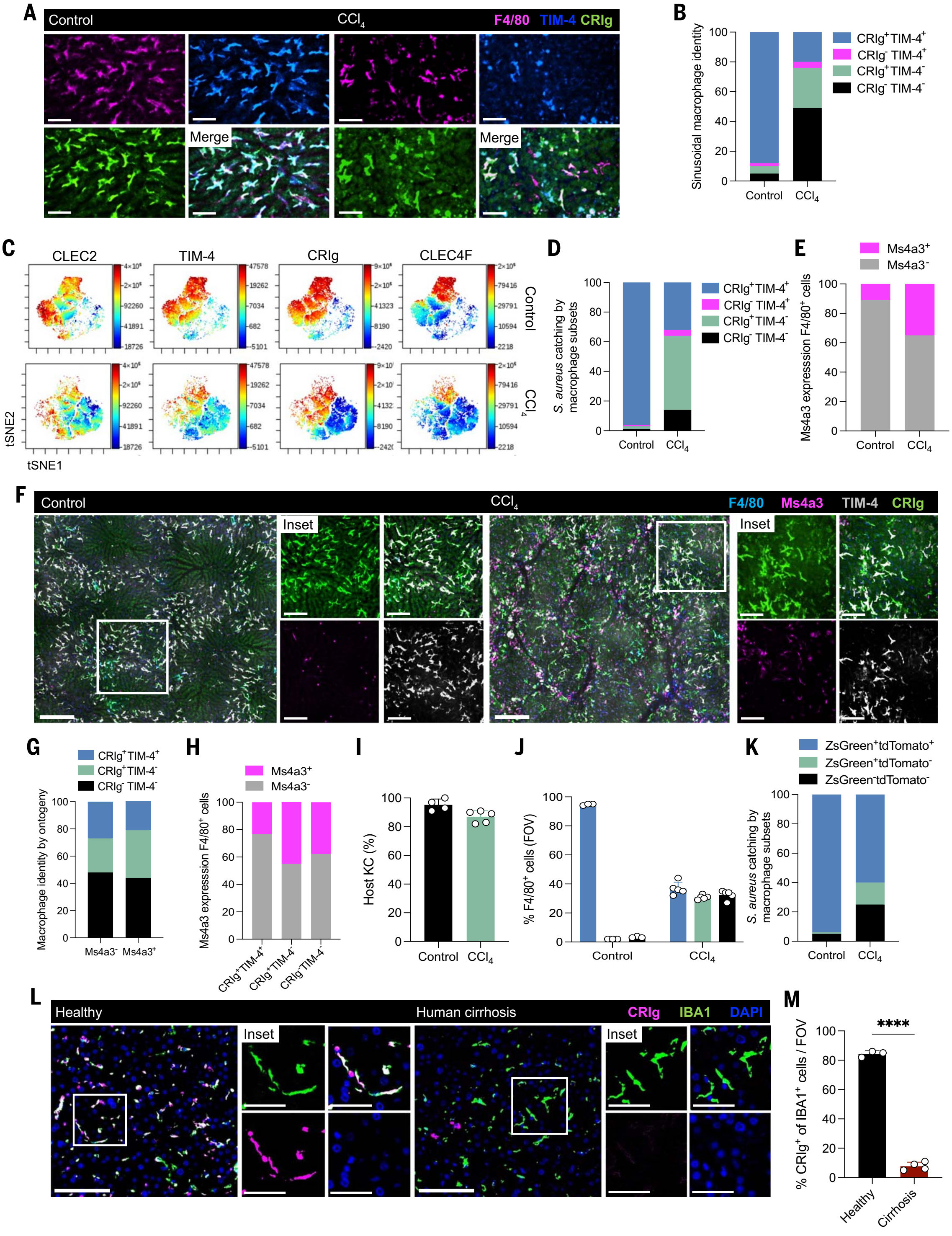


Fig. 3. Liver fibrosis leads to loss of Kupffer cell identity. (**A**) Representative IVM images of KC markers F4/80 (magenta), CRIg (green), and TIM-4 (blue) in control and CCl_4-treated mice. Scale bars: 75 μm. (**B**) Quantification of (A) as a percentage of all F4/80$^+$ cells in sinusoids (n = 4 per group, N = 3). (**C**) tSNE representation showing expression of the key KC markers of CLEC2, TIM-4, CRIg, and CLEC4F in hepatic macrophages. Cells were pre-gated as live singlet CD45$^+$CD19$^-$SiglecF$^-$Ly6G$^-$F4/80$^+$CD11b^{lo}. (**D**) Distribution of *S. aureus* uptake by different liver macrophages by IVM in sinusoids (n = 4 per group, N = 3 or 4). (**E**) Proportion of Ms4a3$^+$F4/80$^+$ cells in sinusoids in healthy and fibrotic mice (n = 3 or 4 per group, N = 2). (**F**) Representative IVM images of Ms4a3$^+$ cells (magenta) in control and fibrotic mice in combination with CRIg (green), TIM-4 (gray), and F4/80 (blue). Scale bars: 250 μm and 100 μm (insets). (**G** and **H**) BM-derived (Ms4a3$^+$) and embryonic (Ms4a3$^-$) hepatic macrophage expression of different KC markers and subsets stratified by Ms4a3 expression assessed by IVM in fibrotic mice (n = 4 per group, N = 2 or 3). (**I**) Quantification of CD45.1 and CD45.2 expression by sinusoidal macrophages in control and CCl_4-treated parabiotic mice (n = 4 or 5 per group, N = 4). (**J**) Quantification of tdTomato, ZsGreen, and F4/80 expression by sinusoidal macrophages assessed by IVM (n = 3 to 5 per group, N = 3). (**K**) Distribution of caught *S. aureus* in the different macrophage subsets in control and CCl_4-treated mice (n = 3 per group, N = 2). (**L**) Representative immunofluorescence images showing KCs in healthy (left) and cirrhotic (right) human livers. Depicted are IBA1$^+$ cells (green) with CRIg (magenta) and nuclei (blue). Scale bars: 120 μm and 50 μm (insets). (**M**) Frequency of CRIg expression by IBA1$^+$ macrophages in healthy and cirrhotic livers (3 or 4 samples per group). Data represented as individual values with mean ± SD. Mann-Whitney test was used for (I), (J), and (M). ****P < 0.0001.

(membrane-spanning 4-domains subfamily A member 3) in peripheral blood was consistently around 95% in both control and fibrotic mice (fig. S2, F and G), as previously reported (22).

Stratifying BM-derived Ms4a3$^+$F4/80$^+$ and embryonic Ms4a3$^-$F4/80$^+$ cells into those expressing CRIg and TIM-4 revealed similar proportions of these markers in both embryonic and BM-derived cells, suggesting that the KCs were de-differentiating during fibrosis (Fig. 3G). In addition, F4/80$^+$CRIg$^+$TIM-4$^+$ cells had the lowest proportion of Ms4a3, but F4/80$^+$CRIg$^+$TIM-4$^-$ and F4/80$^+$CRIg$^-$TIM-4$^-$ cells had a <50% rate of replacement by monocytes (Fig. 3H). Harvesting

all macrophages from liver, not just KCs, revealed that 40 to 50% of F4/80$^+$ hepatic macrophages were Ms4a3$^-$ (fig. S2H), including monocyte recruitment to other liver compartments (subcapsular, etc.). KCs from CCl_4-treated mice had similar levels of CRIg and TIM-4 expression in BM and embryonic macrophages (fig. S2, I and J). We concluded that the lack of niche signals drove tissue-resident and monocyte-derived macrophages asymptotically toward a similar fibrosis-associated phenotype. Parabiosis of CCl_4-treated mice (fig. S2K) and imaging revealed that most F4/80$^+$ cells in liver sinusoids (KCs) were host-derived macrophages, and partner-derived macrophages accounted for <10% at 8 weeks of disease (Fig. 3I and fig. S2L). Moreover, chimerism of host- and partner-derived KCs (F4/80^{hi}CD11b^{lo} cells) was between 5 and 10% (fig. S2, M and N). Similarly, the proportion of KCs expressing CRIg was similar in host- and partner-derived KCs (fig. S2O).

We next used the *Clec4f*$^{cre\text{-}nTdTomato}$×*Rosa26*^{ZsGreen} mouse, a dual reporter for both active expression (tdTomato) and previous expression (ZsGreen) of CLEC4F. More than 95% of F4/80$^+$ cells in sinusoids coexpressed tdTomato and ZsGreen, indicating an active KC phenotype (Fig. 3J and fig. S3A). In CCl_4-treated mice, we found three populations of F4/80$^+$ cells in the sinusoids: tdTomato$^+$ZsGreen$^+$ KCs, tdTomato$^-$ZsGreen$^+$ former KCs (exKCs), and F4/80$^+$ cells with no history of CLEC4F expression (Fig. 3J and fig. S3A). This suggested that some KCs had down-regulated CLEC4F in addition to the KC markers CRIg and/or TIM-4 (fig. S3B).

Principal components analysis (PCA) of sorted F4/80^{hi}CD11b^{lo}Ms4a3$^-$ embryonic KCs from both control and CCl_4-treated mice revealed distinct gene expression profiles (fig. S3C). Out of 22,109 genes, embryonic liver macrophages from fibrotic livers differentially expressed 4298 genes compared with healthy controls, including down-regulation of *Clec4f*, *Vsig4* (*CRIg*), and *Timd4* (fig. S3D). Recent studies have identified that key transcription factors control KC identity, (*23–25*) and are expressed as a result of direct contact with parenchymal cells (*4*). LXRα (*Nr1h3*) is dependent on contact with HSCs and LSECs, whereas ID3 requires hepatocytes (*4*). In agreement with this model, KCs down-regulated *Nr1h3* as they lost contact with HSCs (fig. S3D). By contrast, *Id3* levels did not change, possibly because hepatocyte-derived ID3-activating and/or soluble molecules could still reach the KCs in the context of fibrosis.

We challenged *Ms4a3*^{cre}×*Rosa*^{tdtomato} mice with *S. aureus* MW2 and performed flow cytometric analyses 20 min later (fig. S3, F to H). Although the frequency of F4/80$^+$ cells containing *S. aureus* MW2 was reduced in CCl_4-treated mice, there were no detectable differences between Ms4a3$^+$ and Ms4a3$^-$ cells (fig. S3E). Furthermore, CRIg$^+$TIM-4$^+$ cells (Ms4a3$^+$ and Ms4a3$^-$) caught the most bacteria, irrespective of ontogeny (fig. S3, F and G). In *Clec4f*$^{cre\text{-}nTdTomato}$×*Rosa26*^{ZsGreen} mice treated with CCl_4, bacterial uptake was superior in ZsGreen$^+$tdTomato$^+$F4/80$^+$ KCs (Fig. 3K and fig. S3H), demonstrating that a KC phenotype was required for proficient bacterial capture even in fibrosis. Loss of KC identity could also be seen in liver biopsies of patients with liver cirrhosis. There was a loss of CRIg in IBA1$^+$ human liver macrophages in sinusoids compared with healthy control samples (Fig. 3, L and M).

Physical factors such as reduced blood flow also affected the capture ability of the KCs. Reducing portal flow by 70% in healthy mice (fig. S3I) reduced bacterial capture (fig. S3J). Bacteria still circulated through the sinusoids but were not caught, as evidenced by a sharp rise in free-flowing bacteria during reduced flow (fig. S3J). Thus, a certain level of shear was required for adequate bacterial capture, which was lost in fibrotic sinusoids.

Kupffer cell–like macrophage clusters form with enhanced bacterial capture proficiency

Despite significantly impaired KC capture and killing ability in the sinusoids of fibrotic livers, the majority of bacteria were trapped by this organ (Fig. 2K). Furthermore, there were only small differences in overall survival after infection with 5×10^7 colony-forming units (CFU) of *S. aureus* MW2. Over 7 days, 100% of control mice survived, whereas 75 to 80% survived in the CCl_4-treated group (Fig. 4A). By contrast, depletion of macrophages in control mice using clodronate liposomes (fig. S4A) or genetically depleting KCs (*Clec4f*$^{cre\text{-}nTdTomato}$×*Rosa26*^{iDTR} mouse) (Fig. 4B and fig. S4B) induced 100 and 80% mortality, respectively, within 48 hours after *S. aureus* MW2 infection. In healthy mice, KCs were localized exclusively in liver sinusoids and not larger venules (fig. S4C). By contrast, fibrotic mice showed progressive accumulation of F4/80$^+$ cells in larger venules (fig. S4C). By 8 weeks, CCl_4-treated mice formed large aggregates of F4/80$^+$ macrophages that occupied most of the width of large collateral vessel lumina (Fig. 4C). Collateral vessels appeared to have an abundance of these clusters in larger stitched images, especially at bifurcations, with the occasional individual F4/80$^+$ macrophage observed within these larger vessels (Fig. 4C).

Upon challenge with intravenous *S. aureus* MW2, these macrophage aggregates were extraordinarily efficient at capturing bacteria out of the bloodstream of the collaterals under high shear conditions (Movie 1 and movie S7). In control mice, infection with 5×10^7 *S. aureus* MW2 resulted in an evenly distributed uptake of bacteria by KCs in the liver sinusoids, with 1 to 2 staphylococci captured per KC (movies S2 and S7). By contrast, in fibrotic livers, individual clusters captured large numbers of bacteria (Movie 1, fig. S4D, and movie S7), in some instances exceeding 20 to 30 staphylococci per cluster. Moreover, most large clusters expressed high levels of CRIg (Fig. 4D and movie S7). *CRIg*$^{-/-}$ mice also formed macrophage aggregates after 8 weeks of CCl_4 treatment but could not capture bacteria (Fig. 4G and fig. S4E) and had increased bacteremia and reduced CFU in the liver after intravenous administration of *S. aureus* MW2 compared with wild-type (WT) CCl_4-treated mice (Fig. 4H). These clusters also rapidly acidified pHrodo *S. aureus* bioparticles (Fig. 4I and fig. S4F) at kinetics similar to KCs in control mice. This very efficient clearance of bacteria from high-flow vessels by macrophage aggregates was drastically reduced by mechanical reduction of hepatic flow and quickly restored after blood flow was normalized (Fig. 4J). Thus, the emergence of CRIg-expressing macrophage clusters shifted the capture capacity of the liver from sinusoids to collateral venules in fibrotic livers.

Multinucleated macrophage syncytia are found in human liver diseases

High-resolution microscopy (Fig. 4K) and transmission electron microscopy (TEM) (Fig. 4L) identified multinucleated giant macrophages in mouse fibrotic livers, indicating that in some instances these cells fused into giant cells. However, our imaging unveiled a spectrum of macrophage phenotypes ranging from clusters of individual cells up to and including multinucleated giant cells. Therefore, we called these structures "Kupffer cell–like syncytia."

In a multicenter approach with three liver transplant centers, we systematically assessed human cirrhotic liver tissue. We found numerous clusters of CD68$^+$ multinucleated giant cells in patients with liver cirrhosis from different etiologies, including chronic viral hepatitis, alcoholic liver disease, nonalcoholic fatty liver disease, and cholestatic liver diseases (Fig. 4, M to O, and fig. S4, G and H). IBA1 staining, which identifies all macrophages (*26*), revealed the presence of intravascularly located large macrophages with multiple nuclei (Fig. 4, M to O, and fig. S4H). These KC-like syncytia were also positive for CRIg (Fig. 4N).

Kupffer cell–like syncytia are of bone marrow origin

With their intravascular location, high surface expression of CRIg, and superior bacterial capture capacity, syncytia resembled bona fide KCs. Indeed, IVM revealed the KC-like syncytia expressed TIM-4 and CLEC4F (Fig. 5, A to D, and fig. S5, A and B). Furthermore, the KC-like syncytia were in close proximity to the HSCs that had relocated around the large collaterals, potentially inducing the KC molecules (fig. S1H). Despite their KC-like phenotype, ~90% of the syncytia expressed Ms4a3 (Fig. 5, C and D, and fig. S5B), consistent with their primarily monocyte-derived origin. Liver injury leads

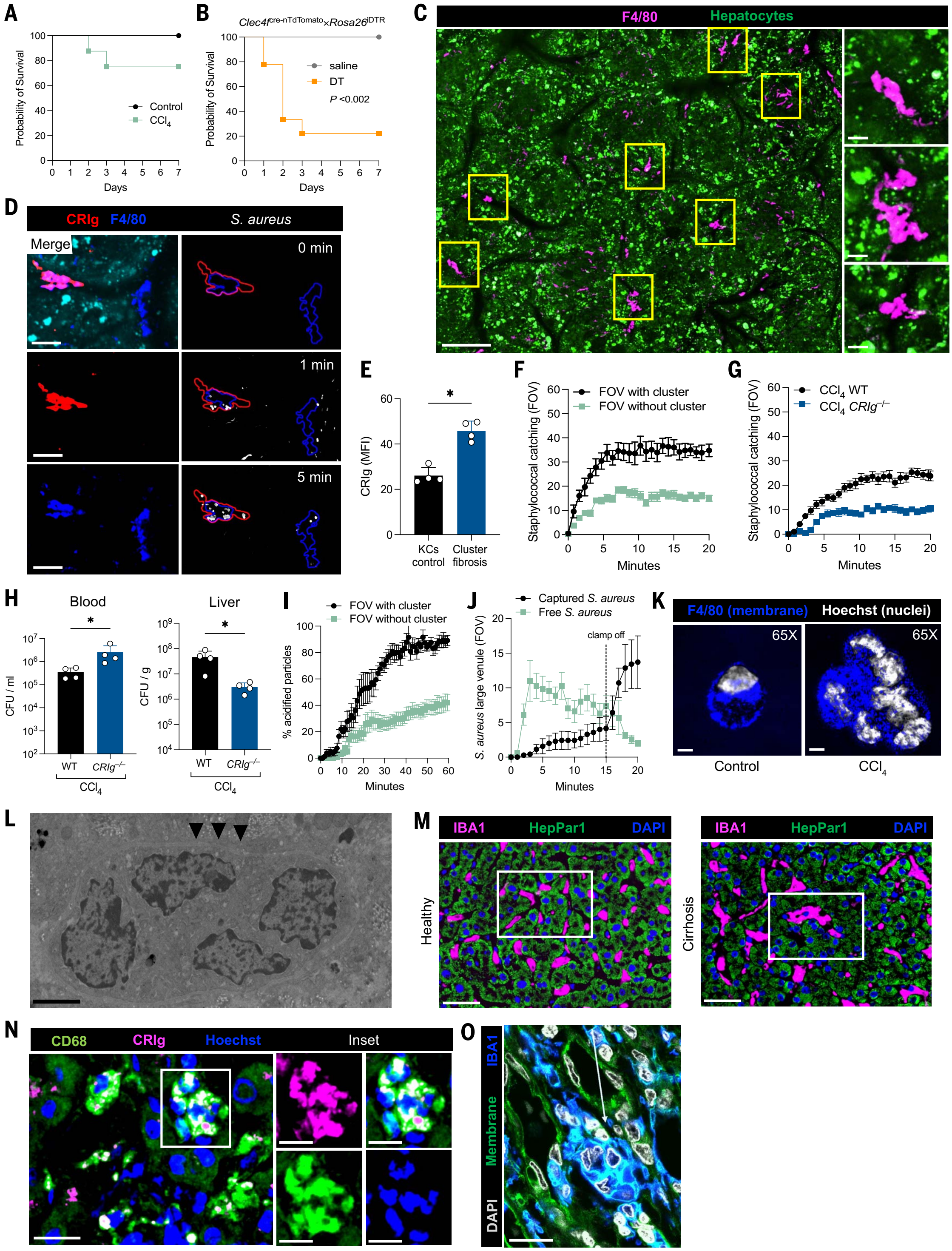
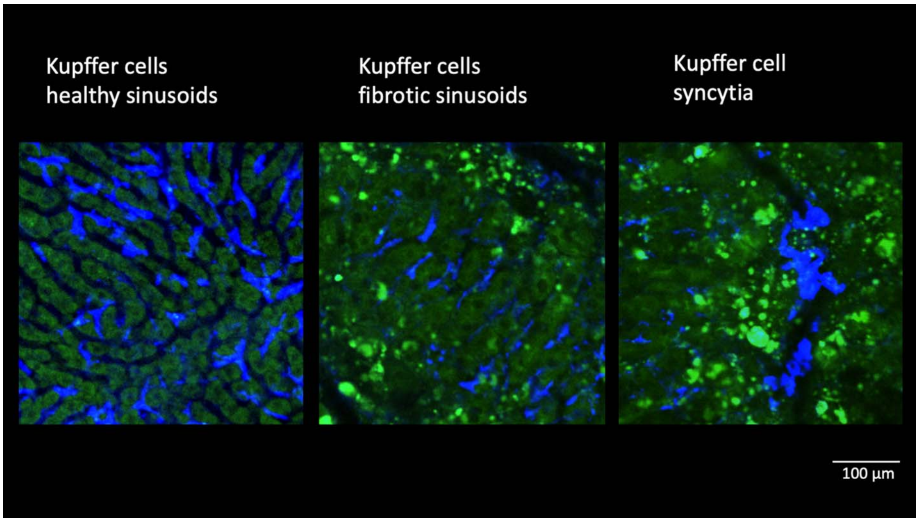


Fig. 4. Multinucleated syncytia emerge in fibrosis with enhanced Kupffer cell features. (**A** and **B**) Survival of mice infected with 5×10^7 CFU *S. aureus* MW2. (A) Control and CCl$_4$-treated mice (n = 8 per group, N = 2) and (B) *Clec4f*$^{cre\text{-}nTdTomato}$ × *Rosa26*^{iDTR} mice 24 hours after application of diphtheria toxin or saline (n = 7 to 9 per group, N = 2). (**C**) Representative stitched IVM showing clusters of F4/80^{+} macrophages (magenta) in large vessels (inset), hepatocytes (dull green), and tissue autofluorescence (bright green). Scale bars: 250 μm and 20 μm (insets). (**D**) Macrophage cluster (blue) at vascular bifurcation expressing high levels of CRIg (red). Scale bars: 50 μm. (**E**) CRIg mean fluorescent intensity of KCs in control mice and macrophage clusters in fibrotic mice (n = 4 per group, N = 2). (**F**) Staphylococcal capture in CCl$_4$-treated mice comparing FOV with and without clusters in large vessels (n = 3 or 4 per group, N = 4). (**G**) Quantification of staphylococcal capture in WT or *CRIg*$^{-/-}$ fibrotic mice (n = 4 per group, N = 2). (**H**) Bacterial burden in WT and *CRIg*$^{-/-}$ fibrotic mice infected with 5×10^7 CFU *S. aureus* (n = 4 per group, N = 2). (**I**) Kinetics of particle acidification by macrophage clusters (n = 3 per group, N = 2). (**J**) Quantification of staphylococcal capture in fibrotic venules with the portal vein clamped and removed after 15 min. (**K**) Ex vivo isolated macrophages from control and fibrotic livers. F4/80 (blue) and Hoechst dye (white) indicate plasma membrane and nuclei, respectively. Scale bars: 10 μm. (**L**) Electron microscopy image of multinucleated giant cell in a CCl$_4$-treated mouse. Arrows indicate the cell wall. Scale bar: 10 μm. (**M**) Representative immunofluorescent staining with IBA1 (macrophages, magenta), HepPar1 (hepatocytes, green), and DAPI (nuclei, blue) showing multinucleated intravascular macrophages in cirrhosis. Scale bars: 50 μm. (**N**) Representative immunofluorescence image of multinucleated macrophages in cirrhosis expressing CD68 (green), CRIg (magenta), and nuclei (blue) shown. Scale bars: 50 μm and 25 μm (insets). (**O**) Multinucleated syncytia in liver cirrhosis showing IBA1 (blue), DAPI (white), and sodium potassium ATPase (green) to label cell membrane. Scale bar: 30 μm. Data represented as individual values with mean ± SD and as mean ± SEM in (F), (G), (I), and (J). Mann-Whitney test was used in (E) and (H). Log-rank test was used for (A) and (B). *P < 0.05, **P < 0.01.

to prominent recruitment of monocytes (27). IVM revealed extensive recruitment of CX3CR1-expressing monocytes within the larger vessels but not sinusoids at 2 weeks of CCl$_4$ treatment, with peak levels at 4 weeks (Fig. 5E and fig. S5C). CX3CR1-expressing monocytes were located in large vessels and started to express F4/80 (fig. S5C), whereas KCs, which lack *Cx3cr1* (28), were positioned in the surrounding sinusoids (Fig. 5E). Imaging of parabiotic mice whose circulatory systems were conjoined revealed that the KC-like syncytia comprised both



Movie 1. The KC capture function in health and disease. Part 1: Bacterial capture of injected *S. aureus* (white) in control mice is performed rapidly by KCs in the sinusoids (blue). Part 2: Bacterial capture in fibrotic sinusoids is suboptimal, with fewer bacteria caught (white) by morphologically altered KCs (blue). Part 3: Rescue KC-like syncytia occupy larger venules in fibrotic livers and catch many bacteria. Movie runtime: 15 min. Scale bar: 100 μm.

host (CD45.2) and partner-derived (CD45.1) cells. Thus, we conclude that circulating monocytes contribute to the formation of KC-like syncytia (fig. S5D).

Monocytes recruited at 4 weeks of CCl_4 treatment but not earlier were the source of KC-like syncytia (Fig. 5, F and G, and fig. S5E), as revealed by tamoxifen-inducible fate mapping using *Cx3cr1*$^{\text{YFP-creER}}$×*Rosa26*$^{\text{TdTomato}}$ mice (*29*). PKH dye was next administered to mice before induction of CCl_4 for long-term labeling of resident macrophages (*30*). Sinusoidal KCs but not the syncytia were labeled with PKH after 8 weeks of CCl_4 treatment, suggesting that KCs did not contribute to syncytia (Fig. 5, H and I). Notably, these KC-like syncytia only formed within the fibrotic liver and not in other organs, such as the lungs or spleen (fig. S5, F and G).

The microbiome drives monocyte recruitment and syncytial formation through MyD88 signaling

Bacterial translocation from the gut is a feature of human cirrhosis (*31*). Mice treated with a cocktail of broad-spectrum antibiotics from birth (Fig. 6A) or germ-free (GF) mice exhibited greatly reduced syncytial formation after CCl_4 treatment (Fig. 6B). Moreover, intravenous administration of *S. aureus* MW2 resulted in reduced bacterial capture by the liver and increased bacterial numbers in blood of antibiotics-treated or GF mice (Fig. 6, C to F). The lack of a microbiome had no significant effect on the development of CCl_4-driven fibrosis, however (fig. S6A). *Myd88*$^{-/-}$ mice treated with CCl_4 also exhibited reduced syncytial formation (Fig. 6G and fig. S6B), impaired bacterial capture, and increased bacteremia (fig. S6C). Thus, the microbiome appears to play a crucial part in the formation of KC-like syncytia.

CD44 is critical in the selective recruitment of monocytes to form Kupffer cell–like syncytia

To better understand KC-like syncytial formation, we interrogated a recently published transcriptomic dataset of mononuclear phagocytes in the CCl_4 fibrosis model and human liver cirrhosis (*12*). Using the same clustering strategy (fig. S6D) and focusing on KC and monocyte gene expression for cell fusion and adhesion molecules listed in the Gene Expression Omnibus (GEO) term “cell–cell fusion” (GO:0140253), *Cd9*, *Cd36*, and *Cd44* were some of the most highly up-regulated genes associated with adhesion and fusion. By contrast, known molecules involved in macrophage fusion (*32*), including *Dcstamp*, *Il4*, and *Tnfrsf11a* (RANK) were not highly expressed in recruited monocytes or macrophages in the fibrotic liver (fig. S6, E and F).

CD44 is the dominant adhesive mechanism for numerous cell types in the liver (*33*–*36*). The large increase in monocytes in the collateral vessels at 4 weeks of CCl_4 treatment was reduced by >80% in *Cd44*$^{-/-}$ mice (Fig. 6, I and J) leading to no KC-like syncytial formation at 8 weeks after CCl_4 treatment (Fig. 6K and fig. S6G). Because CD44 is expressed by endothelial cells as well as monocytes, we next generated BM chimeras and found that only BM-derived CD44 but not parenchymal cell CD44 was required for KC-like syncytial formation after 8 weeks of CCl_4 treatment (Fig. 6L). Bacterial capture was markedly reduced in *Cd44*$^{-/-}$ mice compared with WT mice treated with CCl_4 (Fig. 6M). *Cd44*$^{-/-}$ mice showed increased bacteremia 30 min after infection and one-third the amount of CFU in fibrotic livers (Fig. 6N). The importance of KC-like syncytia for antimicrobial defense was confirmed in BM chimeras, as mice with KC-like syncytia formation (WT BM → *Cd44*$^{-/-}$ mice) showed superior bacterial capture and reduced bacteremia (Fig. 6, O and P). Only 25% of CCl_4-treated WT mice succumbed to infection, whereas all *Cd44*$^{-/-}$ CCl_4-treated mice died by day 4 after infection (Fig. 6Q). By contrast, all vehicle-treated WT and *Cd44*$^{-/-}$ mice survived infection because, in the absence of liver disease, there was presumably no need for syncytia (Fig. 6Q).

The scavenger receptor CD36 functions as a critical fusion molecule for syncytial generation

The fusion and adhesion molecules *Cd9* and *Cd36* were highly expressed on the monocytes infiltrating fibrotic liver (fig. S6, D to F). Anti-CD9 blocking antibody had no effect on syncytia in mice treated with CCl_4 (fig. S6H). By contrast, *Cd36*$^{-/-}$ mice failed to form syncytia after CCl_4 administration but still had an abundance of $CRIg^+F4/80^+$ cells in the collateral vessels (Fig. 7A), suggesting that *Cd36*$^{-/-}$monocytes were recruited but failed to form syncytia and instead presented as dispersed individual macrophages (Fig. 7A). Individual monocytes took up much less volume within the collaterals compared with KC-like syncytia (Fig. 7, B and C). Because CD36 is expressed by other cells, including hepatocytes, we also generated BM chimeras between WT and *Cd36*$^{-/-}$ mice. Only BM-derived CD36 was required for the formation of KC-like syncytia, however (fig. S7, A and B and Fig. 7D).

Next, we tested bacterial capture in *Cd36*$^{-/-}$ fibrotic mice. CCl_4-treated *Cd36*$^{-/-}$ macrophages lacked the ability to capture bacteria in larger vessels (Fig. 7, E and F, and movie S8) despite expressing ample amounts of cell-surface CRIg. BM chimeras with KC-like syncytia (WT → *Cd36*$^{-/-}$) were more efficient at bacterial capture compared with chimeras without syncytia (*Cd36*$^{-/-}$ → WT) and had enhanced bacterial clearance of the blood stream (Fig. 7, G and H). In addition, *Cd36*$^{-/-}$ mice showed a nearly fourfold increase in bacteremia and concomitantly reduced bacterial counts in the liver (Fig. 7F). This was consistent with the hypothesis that the formation of syncytia is necessary to capture bacteria in large collateral vessels. In *Cd36*$^{-/-}$mice that received no CCl_4, there was no observed impairment in *S. aureus* MW2 capture (fig. S7C), and both WT and *Cd36*$^{-/-}$ control mice survived the 7-day observation period after *S. aureus* MW2 infection. By contrast, CCl_4-treated *Cd36*$^{-/-}$ mice exhibited 80% mortality after *S. aureus* MW2 administration compared with 25% mortality in CCl_4-treated WT mice (Fig. 7I).

Discussion

Two hallmark features of chronic liver diseases are profound architectural changes and the

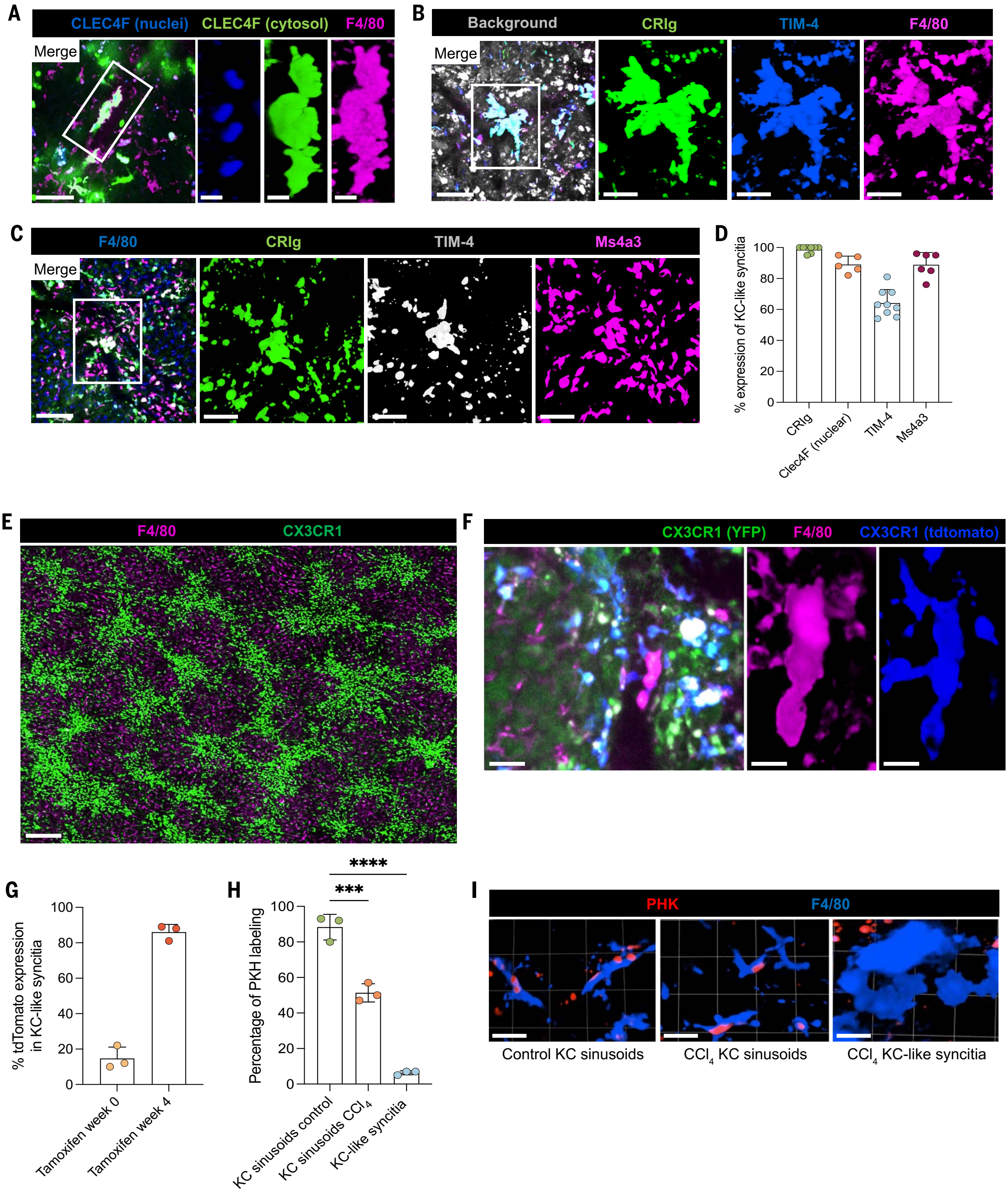


Fig. 5. Syncytia have a Kupffer cell phenotype and are of bone marrow origin. (**A**) 3D-reconstruction of KC-like syncytia in *Clec4f*$^{\text{cre-nTdTomato}}$×*Rosa26*$^{\text{ZsGreen}}$ mice treated with CCl_4. Cytosolic CLEC4F (green) and nuclear CLEC4F (blue) indicate KC phenotype with multiple nuclei together with F4/80 expression (magenta). Scale bars: 100 μm and 20 μm (insets). (**B**) Representative IVM image of KC-like syncytia labeled with TIM-4 (blue), CRIg (green), and F4/80 (purple). Scale bars: 100 μm and 50 μm (insets). (**C**) Representative IVM image of *Ms4a3*$^{\text{cre}}$×*Rosa*$^{\text{tdtomato}}$ mice treated with CCl_4 showing bone marrow origin of KC-like syncytia with Ms4a3 (magenta), CRIg (green), F4/80 (blue), and TIM-4 (gray). Scale bars: 100 μm and 50 μm (insets). (**D**) Quantification of KC marker expression and fate mapping of syncytia. (**E**) Representative stitched IVM image of *Cx3cr1*$^{\text{gfp/wt}}$ mouse treated with CCl_4 (2 weeks) showing accumulation of CX3CR1$^+$ cells (bright green) in pericentral and periportal regions and sinusoidal F4/80$^+$ KCs (magenta). Scale bar: 500 μm. (**F**) Representative IVM with 3D reconstruction of KC-like syncytia labeled with F4/80 (magenta) and fate mapping positive (blue, tdTomato). Hepatocytes in dull green and CX3CR1-expressing cells in bright green (left panel). Scale bars: 50 μm and 30 μm (insets). (**G**) Quantification of tamoxifen-inducible fate mapping at weeks 0 and 4 after tamoxifen injection. (**H** and **I**) Long-term PHK labeling of KCs before CCl_4 treatment. In control mice and in fibrotic sinusoidal KCs, PHK is detected (red). In KC-like syncytia, PKH is absent (n = 3 per group, N = 2). Grids: 25 μm. Data represented as individual values with mean ± SD. One-way ANOVA with Tukey's multiple comparisons test was used for (H). ***P < 0.001, ****P < 0.0001.

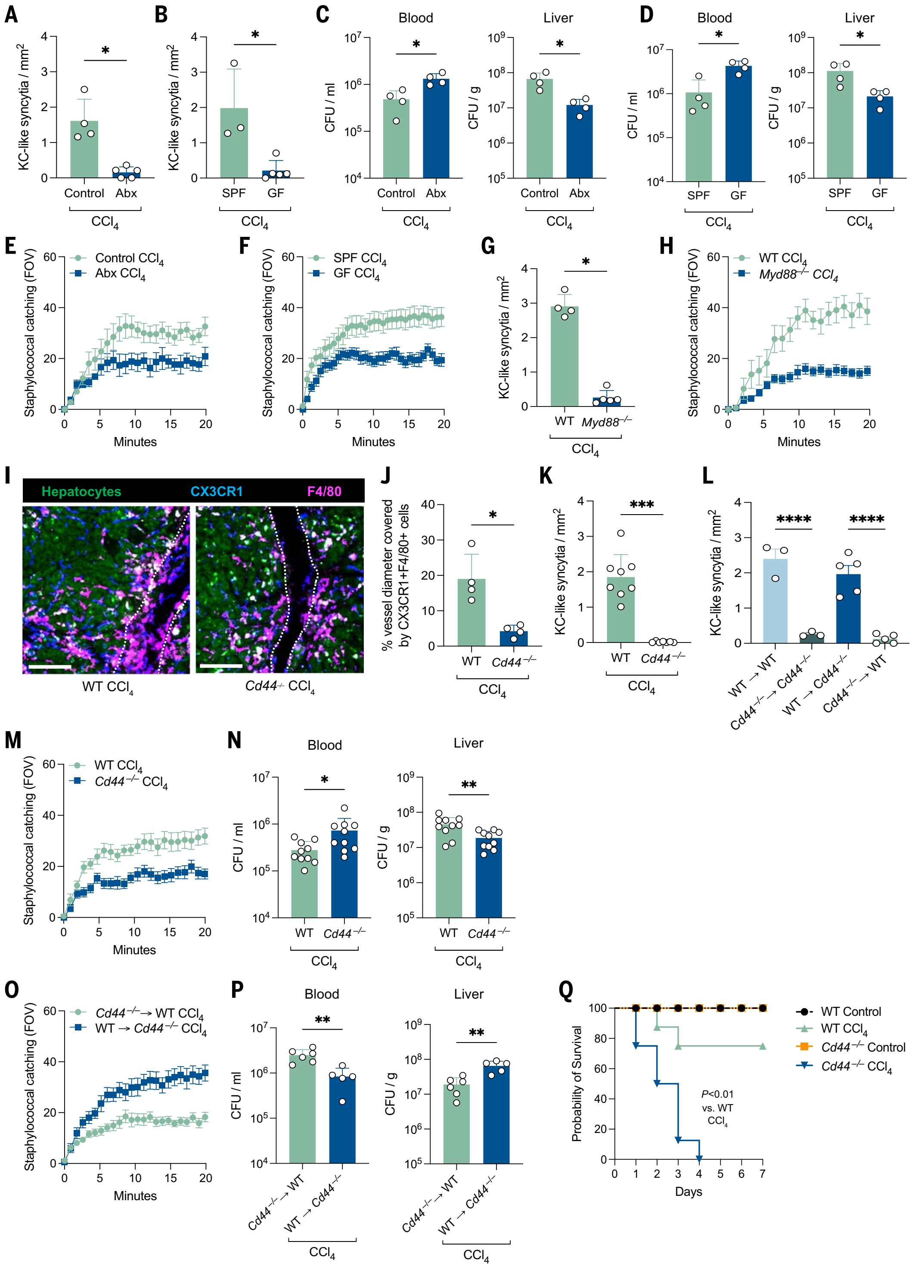


Fig. 6. Kupffer cell–like syncytia are recruited through CD44 and commensal microbes. (**A** and **B**) Quantification of KC-like syncytia in antibiotics-treated and control fibrotic mice (n = 4 or 5 per group, N = 2) and in specific pathogen–free (SPF) and germ-free (GF) fibrotic mice (n = 3 or 4 per group, N = 2). (**C** and **D**) Bacterial dissemination in control and antibiotics-treated (Abx) and SPF and GF fibrotic mice infected with 5×10^7 CFU *S. aureus* MW2 (n = 4 per group, N = 2). (**E** and **F**) Quantification of staphylococcal capture in control and antibiotics-treated and SPF and GF fibrotic mice infected with 5×10^7 CFU *S. aureus* MW2 (n = 4 per group, N = 2). (**G**) Quantification of KC-like syncytia in WT and *Myd88*$^{-/-}$ mice (n = 4 or 5 per group, N = 2) treated with CCl_4. (**H**) Quantification of staphylococcal capture in WT and *Myd88*$^{-/-}$ fibrotic mice (n = 4 per group, N = 2). (**I**) Surface area covered by CX3CR1$^+$ (blue) and F4/80$^+$ (magenta) cells in venules (n = 4 or 5 per group) of WT and *Cd44*$^{-/-}$ mice treated with CCl_4 for 4 weeks. Hepatocytes are depicted in dull green. Scale bars: 100 μm. (**J**) Quantification of (I). (**K** and **L**) Quantification of KC-like syncytia in WT and *Cd44*$^{-/-}$ mice (n = 7 or 8 per group, N = 3) and WT and *Cd44*$^{-/-}$ bone marrow chimeric mice (n = 3 to 5 per group, N = 2). (**M**) Quantification of staphylococcal capture in WT and *Cd44*$^{-/-}$ CCl_4-treated mice (n = 4 per group, N = 2). (**N**) Bacterial burden assessed in WT and *Cd44*$^{-/-}$ CCl_4-treated mice (n = 10 per group, N = 3). (**O**) Quantification of staphylococcal capture in bone marrow chimeras with WT and *Cd44*$^{-/-}$ bone marrow transplantation (n = 4 per group, N = 2). (**P**) Bacterial burden assessed in WT and *Cd44*$^{-/-}$ CCl_4-treated bone marrow chimeras (n = 5 or 6 per group, N = 2) (**Q**) Survival analysis of WT and *Cd44*$^{-/-}$ mice treated with corn oil or CCl_4 for 8 weeks and then infected with 5×10^7 CFU *S. aureus* MW2 (n = 8 per group). Data represented as individual values with mean ± SD, data in (E), (F), (H), (M), and (O) are displayed as mean ± SEM. Mann-Whitney test was used for (A) to (D), (G), (J), (K), (N), and (P). One-way ANOVA with Tukey post hoc test was used in (L). Log-rank test used in (Q). *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

recruitment of monocytes that are thought to augment and replace the pool of KCs found in healthy livers (2, 6, 37). On the basis of imaging and numerous lineage-tracing approaches, we propose that the disappearance of sinusoidal F4/80-expressing macrophages (KCs) is in part due to substantial alteration of their phenotype. They essentially lose both their liver macrophage identity and their primary function, to capture bacteria within sinusoids. These changes are associated with a substantially altered niche. We also find that monocytes enter the liver and acquire the function of KCs. They do not assume the spatial location of KCs, however, as fibrotic sinusoids are poorly perfused, allowing for poor bacterial capture. Our model suggests that these monocytes form KC-like syncytia, shifting antimicrobial defense in the liver from the sinusoids to the large collateral vessels that accommodate most of the blood flow traversing the liver. The increased

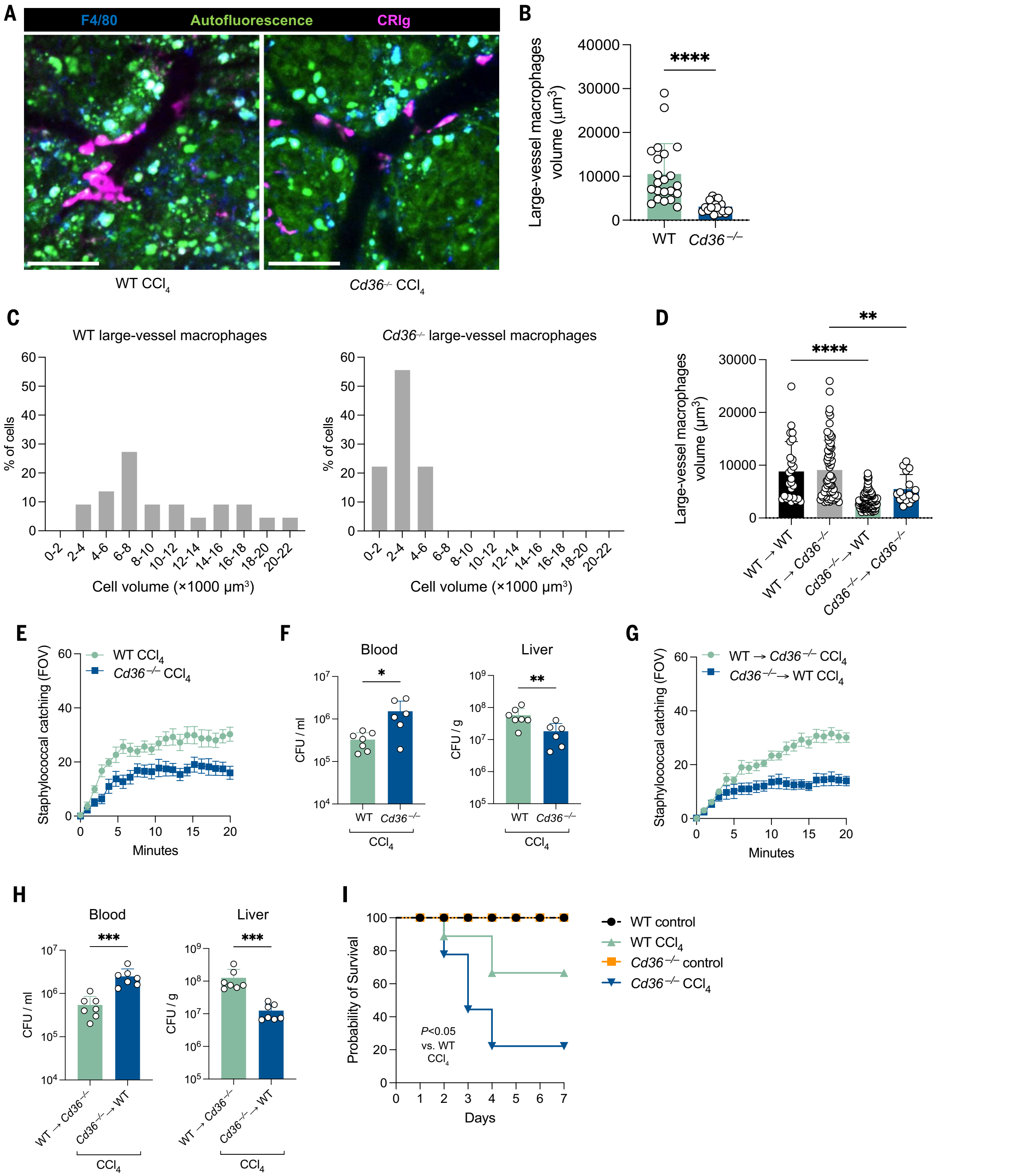


Fig. 7. Kupffer cell–like syncytia depend on the scavenger receptor CD36 for cell adhesion and fusion. (**A**) Representative IVM images of WT and *Cd36*$^{-/-}$ mice treated with CCl_4 for 8 weeks. F4/80 (blue), CRIg (magenta), hepatocytes (dull green), and tissue autofluorescence (bright green) are depicted. Scale bars: 75 µm. (**B** and **C**) Volumetric analysis of CRIg-expressing macrophages in large vessels in WT and *Cd36*$^{-/-}$ fibrotic mice (n = 4 per group, N = 2). (**D**) Volumetric analysis of CRIg-expressing macrophages in large vessels in WT and *Cd36*$^{-/-}$ bone marrow chimeric fibrotic mice (n = 3 to 5 per group, N = 3). (**E**) Enumeration of staphylococcal capture in WT and *Cd36*$^{-/-}$ fibrotic mice infected with 5 × 10^7 CFU *S. aureus* MW2 (n = 5 per group, N = 3). (**F**) Bacterial loads in WT and *Cd36*$^{-/-}$ fibrotic mice infected with 5 × 10^7 CFU *S. aureus* MW2 (n = 6 or 7 per group, N = 2). (**G**) Quantification of staphylococcal capture in WT and *Cd36*$^{-/-}$ bone marrow transplant experiments (n = 3 or 4 per group, N = 2). (**H**) Bacterial loads in WT and *Cd36*$^{-/-}$ bone marrow transplant recipients (n = 7 per group, N = 3). (**I**) Survival analysis of WT and *Cd36*$^{-/-}$ mice treated with CCl_4 for 8 weeks or controls and then infected with 5 × 10^7 CFU *S. aureus* MW2 (n = 6 to 8 per group, N = 2). Individual dots each represent one mouse in (F) and (H) and mean ± SEM in (E) and (G). Mann-Whitney test used for (B), (F), and (H). One-way ANOVA with Tukey post hoc test used in (D). Log-rank test used for survival study in (I). *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

presence of microbes or microbial products in the liver as well as high shear in these vessels cause monocytes to be recruited by CD44. They then fuse into syncytia mediated by CD36 and occupy a substantial part of the vessel lumen, a feature paramount for the capture of bacteria under flow conditions that have been altered by fibrosis. Moreover, a very profound relocation of stellate cells around the enlarged vessels perhaps helps impart KC characteristics to the KC-like syncytia, often capturing 10 times the number of bacteria that a single KC captures in its normal sinusoidal environment while maintaining tremendous microbicidal capacity.

Multinucleated giant cells were first described by Langerhans in tuberculoid granulomas in 1868 (*38*) and have subsequently been identified in a number of pathologic settings, including foreign body responses, atherosclerosis, and autoimmunity (*39*). Giant cells in the bone form membrane fusions thought to increase their phagocytic capacity (*32*). In the liver, KCs have the critical task of capturing particulate matter under flow conditions, which notably precedes the process of phagocytosis. In larger vessels with much greater diameters, this task becomes insurmountable for a single macrophage, as was seen in *Cd36*$^{-/-}$ mice. However, by increasing their size through bridging of multiple cells and filling a greater proportion of the lumen, these KC-like syncytia form a critical intravascular shield to help capture and eradicate bacteria in high-flow vessels. These cells presented as a continuum, ranging from clear clusters of abutting cells to true giant cells with a single membrane. Notably, KC-like syncytia did not block blood flow in a manner analogous to how KCs allow adequate flow in healthy sinusoids despite occupying a considerable proportion of the lumen.

Previous work on the KC niche determined that infiltrating monocytes have to contact LSECs, hepatocytes, and stellate cells, which then activate specific transcription factors to imprint KC identity (*4*, *24*, *40*). Within this framework, parenchymal cells provide instructive signals for KC development to fulfill their tissue-specific functions. We extend this work into disease states by showing that KCs in the context of fibrosis retract their many pseudopods and lose their instructive niche as sinusoids undergo fibrotic remodeling. These KCs downregulate important transcription factors including LXRα (*Nr1h3*) and downstream molecules such as CRIg and TIM-4. Previous single-cell RNA sequencing (RNA-seq) studies are consistent with these discoveries showing the increased heterogeneity of liver macrophages (*12*, *19*).

Currently, 1.5 billion people worldwide live with chronic liver disease, and the burden is rising rapidly (*41*). Despite the importance of the liver as a microbial filter, most of these individuals do not suffer from bacteremia (*42*). Similarly, most fibrotic mice challenged with a large intravenous bolus of pathogen survive infection, suggesting that a conserved presence of CRIg-expressing KC-like syncytia in fibrosis in both humans (*43*, *44*) and rodents may be an adaptive mechanism to protect the host from infections. Although it is difficult to make the argument that alcoholic liver disease and nonalcoholic fatty liver disease have evolutionarily driven these adaptations for survival, liver trophic viruses that infect mammals may have spurred very similar evolutionary adaptations. Thus, the architectural reorganization of the liver disrupts the core constituents of the KC niche. This, in turn, drives an immune reorganization that may be a critical part of mammalian evolution by which the liver responds to severe chronic insults (viral, toxic, etc.) and prolongs organismal survival.

Materials and methods

Mice

All experiments were carried out using 8- to 12-week-old male and female mice. The experimental procedures were approved by the University of Calgary Animal Care Committee and were in accordance with Canadian Council for Animal Care Guidelines (Protocol No. AC-19-0138). All mice were cohoused and bred in a specific pathogen-free facility at the University of Calgary with a 12-hour light–dark cycle and access to food and water ad libitum. C56BL/6 mice (WT) were obtained from the Jackson Laboratory and subsequently bred in house. Pep BoyJ (B6.SJL-*Ptprc*^{a}*Pepc*^{b}/BoyJ, Jax: 002014), *Cx3cr1*^{creER} [B6.129P2(Cg)-*Cx3cr1*$^{tm2.1(cre/ERT2)Litt}$/WganJ, Jax: 021160], Ai9 [B6.Cg-*Gt(ROSA)26Sor*$^{tm9(CAG-tdTomato)Hze}$/J, Jax: 007909], *Cd44*$^{-/-}$ [B6.129(Cg)-*Cd44*^{tm1Hbg}/J, Jax: 005085], *Cd36*$^{-/-}$ (B6.129S1-*Cd36*^{tm1Mfe}/J, Jax: 019006), *Cx3cr1*$^{gfp/+}$ [B6.129P2(Cg)-*Cx3cr1*^{tm1Litt}/J, Jax: 005582], Ai6 [B6.Cg-*Gt(ROSA)26Sor*$^{tm6(CAG-ZsGreen1)Hze}$/J, Jax: 007906], *Rosa26*^{iDTR} [C57BL/6-G*t(ROSA)26Sor*$^{tm1(HBEGF)Awai}$/J, Jax: 007900], and *Myd88*$^{-/-}$ [B6.129P2(SJL)-*Myd88*$^{tm1.1Defr}$/J, Jax: 009088] mice were purchased from the Jackson Laboratory and bred in house. *Clec4f*$^{cre-nTdTomato}$ mice (MGI Ref. ID: J:279524) were a kind gift from C. Glass (University of California, San Diego) and were crossed with Ai6 [B6.Cg-*Gt(ROSA)26Sor*$^{tm6(CAG-ZsGreen1)Hze}$/J] mice to receive the dual reporter and with *Rosa26*^{iDTR} mice for specific ablation. *CRIg*$^{-/-}$ mice (MGI Ref. ID: J:138691) were a kind gift from M. van Lookeren Campagne (Genentech). *Lrat*^{cre} mice (MGI Ref. ID: J:205330) were a kind gift from R. Schwabe (Columbia University, New York), and *Ms4a3*^{cre} mice (MGI Ref. ID: 284702) were provided by F.G. Both were crossed with Ai14 [B6.Cg-Gt(ROSA)*26Sor*$^{tm14(CAG-tdTomato)Hze}$/J] mice.

Antibodies and reagents

All antibodies used in this study for IVM and flow cytometry are summarized in table S1. Additional reagents used for IVM included labeling macrophages with the PKH26 Phagocytic Cell Labeling kit (Sigma-Aldrich) for long-term labeling of KCs, which was prepared according to the manufacturer's instructions. A dose of 100 μM was injected intravenously to label KCs. Tetramethyl rhodamine isothiocyanate (TRITC)–labeled Dextran (Invitrogen, Thermo Fisher Scientific) was used to visualize hepatic blood flow. Vybrant DiD Cell-Labeling Solution, pHrodo Red *Staphylococcus aureus* BioParticles, and Alexa Fluor 647 NHS Ester were obtained from Thermo Fisher Scientific. Tamoxifen and carbon tetrachloride were obtained from Sigma-Aldrich. Anti-CD9 antibody (clone ALB 6) was purchased from Santa Cruz Bioscience, isotype mouse IgG1 from BD Bioscience. Clodronate and phosphate-buffered saline (PBS) liposomes were obtained from https://clodronateliposomes.com/.

Bacterial strains and culture

S. aureus strain MW2 was obtained from the Network on Antimicrobial Resistance in *S. aureus* (NARSA) and transformed with pCM29 (*45*) to constitutively express green or yellow fluorescent protein (*46*). The MW2 strain was used for all experiments. Bacteria were cultured overnight in brain heart infusion medium (Difco) at 37°C on a shaker. Chloramphenicol (10 μg/ml) was added to the medium to maintain the plasmids. Bacteria were subcultured without antibiotics in brain heart infusion media until exponential phase growth was achieved (OD_{660nm} = 1.0), washed with saline, and resuspended in saline.

Mouse treatments

To induce liver fibrosis, mice were treated with 0.75 μl per gram of body weight of carbon tetrachloride (Sigma-Aldrich) dissolved in corn oil (Sigma-Aldrich) or corn oil alone by means of oral gavage three times a week for 8 weeks, unless otherwise specified. Mice were sacrificed 48 hours after the last dose. For infection experiments, mice were infected by tail vein injection of 5×10^7 CFU *S. aureus* MW2 in 200 μl of saline. For survival studies, mice were infected with an intravenous injection of 5×10^7 CFU *S. aureus* MW2 in 200 μl of saline. Infected mice were monitored closely, and their health status was assessed using an established scoring system (*47*). On the basis of this scoring system, mice were euthanized in cases where they reached their humane endpoint. KC depletion was performed as previously described (*14*). We injected 200 μl of clodronate liposomes (690 mM) via the tail vein 48 hours before the experiment or 200 μl of PBS liposomes for control groups. KCs in *Clec4f*$^{cre-nTdTomato}$×*Rosa26*^{iDTR} mice were depleted by a single application of 10 ng per gram of body weight of diphtheria toxin (Sigma-Aldrich), as previously described (*24*). Anti-CD9 antibody (clone ALB 6) and IgG1

isotype were injected at a dose of 10 μg per mouse intraperitoneally three times per week for 8 weeks. Tamoxifen (Sigma-Aldrich) was dissolved in corn oil (Sigma-Aldrich) at 20 mg/ml and administered via intraperitoneal injection (100 μg per gram of body weight) for five consecutive days during indicated timepoints of CCl_4 administration.

Bacteriological analysis

Bacteriological analyses were performed after imaging experiments, 30 min after intravenous injection of bacteria, unless otherwise indicated. Anesthetized mice were washed and disinfected with 70% ethanol. Blood was collected by cardiac puncture into a heparinized syringe (100 U/ml). A liver lobe, spleen, lung lobe, and kidney were carefully removed, weighed, and homogenized using a tissue homogenizer. To assess colony-forming units, we serially diluted blood or tissue homogenate and then plated 30 μl onto brain heart infusion agar plates followed by incubation at 37°C. After 18 hours, bacterial colonies were counted.

Parabiosis

Parabiotic pairs were generated as previously described (*48*). Briefly, age-matched female mice were cohoused for 3 weeks before parabiosis surgery. For surgery, mice were anesthetized with an intraperitoneal injection of ketamine (200 μg per gram of body weight; Bayer Animal Health) and xylazine (10 μg per gram of body weight; Bimeda-MTC) and placed under a sterile hood on a heating pad. A skin incision was made from shoulder levels to hips along lateral flanks on opposite sides. First, sutures were placed on shoulder and thigh skin and muscle layers and subsequently, sutures were completed along the incision of skin flaps using a stapler. Blood chimerism was assessed after 21 days of parabiosis by tail-vein canulation and again after 8 weeks of CCl_4 treatment.

Depletion of commensal bacteria

Mouse intestinal commensal bacteria were depleted at birth with a cocktail of broad-spectrum antibiotics, as previously published (*49*). Mice were administered ampicillin (1 mg/ml), vancomycin (0.5 mg/ml), neomycin (1 mg/ml), metronidazole (1 mg/mg), and ciprofloxacin (0.2 mg/ml) via the drinking water. Antibiotics were purchased from Sigma-Aldrich.

Bone marrow transplantation

We performed two sets of experiments using bone marrow chimeras. In a first set of experiments, WT and *Cd44*$^{-/-}$ mice were used, and in the second set, WT and *Cd36*$^{-/-}$ mice were used. Each experiment included BM transfer controls (WT → WT, *Cd44*$^{-/-}$ → *Cd44*$^{-/-}$, and *Cd36*$^{-/-}$ → *Cd36*$^{-/-}$). Bone marrow was isolated from tibias and femurs of donor mice under sterile conditions. Recipient mice received two doses of radiation of 5.25 grays (Gy) with a 3-hour interval between irradiations. Next, 8×10^6 bone marrow cells were transferred via tail-vein injections. Mice were kept for 8 weeks to allow full bone marrow reconstitution.

Labeling of red blood cells

A donor mouse was sacrificed by cardiac puncture, and 500 μl of heparinized whole blood was obtained. Cells were then centrifuged at 400*g* for 10 min at room temperature, and plasma and leukocytes were discarded. RBCs were resuspended in 1 ml of RPMI medium (Thermo Fisher Scientific) supplemented with 2% fetal bovine serum. Cells were stained with 10 μl of DiD (Thermo Fisher Scientific) at room temperature for 10 min and then washed twice.

Mouse surgery for intravital microscopy of the liver

IVM was performed as previously published (*50*, *51*). Mice were anesthetized using ketamine (200 μg per gram of body weight; Bayer Animal Health) and xylazine (10 μg per gram of body weight; Bimeda-MTC) delivered by intraperitoneal injection. To gain intravenous access, a small catheter was inserted into the tail vein of mice. This allowed us to deliver fluorescently conjugated antibodies, bacteria for capture experiments, and additional anesthetics and fluids to maintain adequate hydration status of the mice. A surgical incision was made in the abdomen, and skin and abdominal muscle were partly removed. The liver was mobilized by cutting the attaching ligaments, while care was taken not to touch the liver to avoid tissue damage. Next, the mouse was placed in a lateral position on a heated microscopy stage with an inserted cover glass. The heated plate allowed us to maintain body temperature at 37°C. The liver was subsequently flipped onto the glass coverslip using a fine cotton tip to remove the intestines in a no-touch technique and then covered with wet laboratory tissues for stabilization. The mouse was wrapped in saline-soaked gauze and continuously hydrated by applying saline. During imaging, mice were regularly administered intravenous fluids and additional anesthetics.

Intravital microscopy of the liver

Images were acquired using an inverted spinning-disk confocal microscope (IX81; Olympus) that was equipped with a focus drive by Olympus and a motorized stage (Applied Scientific Instrumentation) to allow live movement in x and y axes. The microscope was stocked with a motorized objective turret fitted with 4X/0.16 UPLANSAPO, 10X/0.40 UPLANSAPO, and 20X/0.70 UPLANSAPO objective lenses. The microscope was placed onto an antivibration customized table (Newport, Irvine, CA). The spinning-disk confocal microscope was linked to a confocal light path (WaveFx; Quorum Technologies) that was based on a modified CSU-10 head (Yokogawa Electric Corporation). Laser excitation wavelengths were 491, 561, and 642 nm (Cobolt, Stockholm, Sweden), and a 512 pixel by 512 pixel back-thinned EMCCD camera (C9100-13, Hamamatsu, Bridgewater, NJ) was used for detection. Laser wavelengths were merged into an optic cable using an LMM5 laser merge module (Spectral Applied Research, Richmond Hill, Canada). Fluorescence visualization was through the bad-pass emission filters (Semrock, Rochester, NY) ET 525/50M, FF 593/40, and ET 700/75M and driven by a MAC 6000 Modular Automation Controller (Ludl Electronic Products, Hawthorne, NY). Volocity software (Quorum) was used to drive the confocal microscope and for acquisition and analysis of images. A combination of fluorescently conjugated antibodies, reporter mice, or fluorescent bacteria were used to visualize cells of interest and as functional readouts. For bacterial capture and imaging of particles, three to five fields of view (FOV) were randomly selected in the sinusoid area and three to five FOV in the venular region with the 10× objective and imaged for 20 min. For erythrocyte tracking, 1-min videos were obtained from a single point with low exposure time to allow for continuous imaging of flow. Collagen was imaged using a Leica SP8 DIVE inverted system with a multiphoton laser (*52*). An HC FLUOTAR L 25X/0.95 water immersion objective was used at 2× digital zoom. The microscope was equipped with a tunable InSight X3 ultrafast laser (Spectra-Physics), excited at 800 to 1000 nm and detected with external, extremely sensitive non-descanned hybrid detectors. Second-harmonic generation was imaged by tuning the detector to the excitation wavelength divided by two. LAS X software was used to drive the microscope.

Mechanical alteration of liver hemodynamics

To manipulate hepatic blood flow, we induced a mechanical stenosis of the portal vein. A small vascular clamp (Fine Surgical Tools, Vancouver, BC, Canada) was used, and a 4-0 Prolene suture (Ethicon) was tied around one end to prevent full occlusion. Mice were prepared for IVM as described. Before the liver was mounted onto the cover glass, the portal vein was identified, and the clamp was carefully applied. In some experiments, as indicated, the clamp was removed during live imaging to restore blood flow.

Intravital microscopy of the lung

Lung IVM was performed as described previously (*30*). Briefly, anesthetized mice were placed on a heating pad. A tracheotomy was performed, and mice were ventilated using a small rodent ventilator (Harvard Apparatus). After the mouse was placed in a lateral position, a small intercostal incision was applied, and a custom-made intercostal lung window was inserted. Lung tissue was stabilized using suction of 20 mmHg.

Images were acquired using an upright microscope (BX51, Olympus) equipped with a confocal light path (WaveFx, Quorum, ON, Canada) and a back-thinned electron multiplying charge-coupled device camera (Hamamatsu Photonics) for fluorescence detection. Lung macrophages were visualized by intratracheal injection of 0.5 μM PKH26 (Sigma Aldrich, Darmstadt, Germany) and intravenous injection of fluorescently labeled anti-CX3CR1 monoclonal antibody (1 μg per mouse; clone SA011F11; Biolegend).

Intravital microscopy of the spleen

Spleen IVM was performed as previously described (*48*). In brief, a skin incision was made, the mouse was placed in a left lateral position, and the spleen was exteriorized onto a cover glass and fixed with wet laboratory tissue paper. A suture was placed in the attached fat tissue to stabilize the spleen on the cover glass. Red-pulp macrophages and vasculature were visualized using fluorescently labeled antibodies against CD31 and F4/80 (table S1). Images were acquired using an inverted spinning disk confocal microscope (IX81; Olympus), which was coupled with a confocal light path (WaveFx, Quorum, ON, Canada), a focus drive, and a motorized stage (Applied Scientific Instrumentation). A back-thinned electron multiplying charge-coupled device camera (Hamamatsu Photonics) was used for fluorescence detection, and Volocity Software (Quorum, ON, Canada) was used to drive the microscope.

In vivo phagocytosis assay

Phagocytosis and lysosome acidification of liver macrophages were assessed as previously described (*15*). Following the manufacturer's instructions, pHrodo Red *S. aureus* BioParticles (2 mg/ml) were stained with 50 μg/ml of Alexa Fluor 647 NHS ester in 100 mM bicarbonate-buffered saline (pH 8.3) for 30 min at room temperature. For IVM, fields of view were chosen, and a baseline was recorded for 1 min before injection of bioparticles (40 μg per mouse). The pHrodo and AF647 signals were visualized for 60 min. Volocity software was used to quantify uptake and percentage of acidified particles. The bioparticles were labeled with AF647 as a reference color, and acidification of the bioparticles in low pH environment in phagolysosomes was indicated by an increase in pHrodo in the caught and internalized bioparticles.

Isolation of hepatic leukocytes

Liver leukocytes were isolated as previously described (*50*). In brief, the inferior vena cava was cannulated, and the liver was perfused in situ, first with Ca^{2+}-, Mg^{2+}-free Hank's Balanced Salt Solution (HBSS, Thermo Fischer Scientific) and then with HBSS containing 0.05% collagenase type IV (Worthington Biomedical, Lakewood, NJ) and 0.02% DNase I (Roche) at a flow rate of 4 ml/min using a peristaltic pump. The liver was subsequently removed and digested in HBSS containing 0.009% collagenase type IV and 0.02% DNase I at 37°C for 30 min on a shaker. The suspension was then passed through a 100-μm cell strainer (BD Bioscience) and centrifuged at 25*g* for 5 min at room temperature to remove debris and hepatocytes. The supernatant was collected and centrifuged at 400*g* for 10 min at 4°C. The cells were then resuspended in 15% iodixanol solution (Sigma-Aldrich) and centrifuged at 400*g* for 15 min at room temperature without a break. The band of nonparenchymal cell fraction was aspirated using a transfer pipet and washed twice with HBSS. After a step of red blood cell lysis using ACK (ammonium-chloride-potassium) lysis buffer (Thermo Fisher Scientific), the cells of interest were resuspended in 100 μl of flow buffer [PBS with 0.5% bovine serum albumin (BSA) and 2 mM EDTA].

Flow cytometry

Resuspended cells were first blocked using anti-mouse CD16/32 antibody (2.4G2, BioXcell) for 10 min. Cells were then stained with specific markers for 30 min at 4°C, and appropriate isotype antibodies and FMO controls were used. Nonviable cells were labeled using viability dye Ghost Dye red 710 (TONBO Bioscience). After staining, cells were washed and fixed in 1% paraformaldehyde (PFA). Spectral multicolor flow cytometry was performed immediately after staining without fixing cells. The samples were acquired using a BD FACS Canto II flow cytometer or a spectral flow cytometer (Aurora, Cytek) and analyzed using FlowJo software (Tree Star). Cells were gated on forward scatter/side scatter, singlets, live, and $CD45^+$ to identify leukocytes. Subsequently, neutrophils were identified as $CD45^+Ly6G^+$ cells and nongranulocytes were analyzed further after exclusion of $SiglecF^+$ cells and $CD19^+$ cells. Cell populations from multiple mice per group (control and CCl_4) were concatenated, and simultaneous tSNE analyses were performed.

Isolation and ex vivo imaging of liver macrophages

Single-cell liver suspensions were generated as described above, except cells were not passed through a filter. Instead, single-cell suspensions were placed in a 15-ml canonical tube and allowed to rest for 5 min at room temperature to sediment debris and hepatocytes. Supernatant was aspirated, and debris and hepatocytes discarded. The supernatant was centrifuged for 10 min at 400*g*. Cells were then resuspended in 15% iodixanol solution and centrifuged at 400*g* for 15 min at room temperature without a break. Cells were resuspended in Ca^{2+}-, Mg^{2+}-free Hank's balanced salt solution (HBSS, Thermo Fischer Scientific). Nonparenchymal liver cells were collected and stained with F4/80 antibodies and Hoechst nuclear dye (Invitrogen, Thermo Fisher Scientific) in a buffer containing 0.5% BSA and 2 mM EDTA. Cells were fixed for 10 min in 4% paraformaldehyde. Single-cell suspensions were mounted on a coverslip and imaged immediately on a Leica SP8 DIVE inverted system using the 65× water immersion objective in high magnification.

Transmission electron microscopy

Mice were sacrificed and the liver was excised. Punch biopsies were obtained (1 mm by 3 mm) and fixed in a solution of 1.6% paraformaldehyde and 2.5% glutaraldehyde in 0.1 M cacodylate buffer. Sections were dehydrated through an ethanol series and embedded in epon mixture. The epon layer with tissue was removed after polymerizing. A representative area was identified using a light microscope and glued to resinstub for sectioning. Fine sections were cut using an ultramicrotome with a diamond knife and collected on single-hole grids with Formvar support. Sections were subsequently stained with aqueous uranyl acetate and Reynold's lead citrate. Images were acquired with a Hitachi H7650 TEM at 80 kV equipped with a AMT16000 digital camera.

Human samples

Human liver tissue was obtained from three major transplant hepatology centers: the Department of Hepatobiliary and Transplant Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Charité University Medicine, Berlin, Germany; and the Multi-Organ Transplant Program, Toronto General Hospital Research Institute, University of Toronto, Toronto, Canada. Healthy tissue was obtained from neurologically deceased donors whose livers were determined to be acceptable for transplantation. Cirrhotic tissue was obtained from explant livers or patients undergoing liver biopsy. The etiology of cirrhosis of the patients included was alcoholic liver disease (n = 4), nonalcoholic fatty liver disease (n = 5), hepatitis B virus infection (n = 1), primary sclerosing cholangitis (n = 5), and primary biliary cholangitis (n = 1). All patients provided written informed consent. Study was approved by the Institutional Review Board of the medical faculty at the University of Hamburg (PV4898), the Ethics Committee at Charité University Medicine Berlin (Project EA2/091/19), and the institutional Research Ethics Board in Toronto (REB #14-7425 and REB # 20-5142) and conducted according to the ethical principles of the Declaration of Helsinki and its revised versions.

Immunohistochemistry of human liver tissue

Formalin-fixed paraffin-embedded (FFPE) liver tissues were sectioned (4-μm sections) and then stained with hematoxylin and eosin (H&E), trichrome, or anti-human CD68 according to

the standard histological procedures. Additional antibodies used for immunohistochemistry were anti-IBA1 (clone 20A12.1, Merck, Germany), anti-sodium potassium adenosine triphosphatase (ATPase) antibody (clone 464.6, Abcam, USA), anti-Heppar1 (clone OCH1E5, Dako, Denmark), and anti-VSIG4 (clone EPR22576-125, Abcam, USA).

Immunofluorescence staining of liver cryosections

Liver tissues were embedded in cryo molds using Shandon Cryochrome (ThermoFisher Scientific), and 5-μm tissue sections were prepared. Tissue slides were fixed and permeabilized in acetone for 10 min at room temperature. Tissue sections were washed three times with PBS and blocked by incubating with 10% BSA/PBS for 30 min at RT. Subsequently, tissue sections were incubated with the primary antibodies CD68-PE (clone: JO127; Santa Cruz; dilution: 1:100) and VSIG4-APC (CRIg) (clone: JAV4, eBioscience; dilution: 1:100) overnight at 4°C. Tissue sections were washed three times with 0.1% BSA/PBS and stained with Hoechst 34580 (Sigma-Aldrich; dilution: 1:2000) for 15 min at RT. Tissue sections were mounted using fluorescein mounting media (Dako, Glostrup, Denmark) and imaged using an Nikon A1 confocal microscope equipped with a 40× NA 1.15 water immersion long-working-distance objective.

Fibrosis assessment

Livers were fixed in 10% formalin and sent to Calgary Laboratory Services for paraffin embedding, sectioning, and staining of Sirius Red. Slides were imaged using a Thorlabs Whole-Slide-Scanning microscope. Sirius Red staining was quantified using ImageJ (53).

Image analysis and processing

Image analysis was done using Volocity Software (Quorum), ImageJ (NIH), or Imaris Software (Bitplane). Immunofluorescence data was analyzed using ImageJ, images were exported as .tif and further analyzed using ImageJ (NIH). Bacterial capture experiments were assessed using Volocity software as previously described (15) in automated fashion using the "find objects" function with the same threshold settings throughout. Background autofluorescence was assessed before injection of bacteria and subtracted from the results. Percentage of acidified particles was calculated from the total number of captured bioparticles with the reference color and the particles that over time turned red. For the measurement of sinusoid and venule diameters, images were exported from Volocity software as .tif files. ImageJ software (NIH) was used to measure diameters. We measured 10 randomly selected sinusoids and venules in five fields of view (FOV) per mouse. Syncytia quantity was performed using computer-generated stitch images exported from Volocity. Liver surface area was determined, and syncytia were counted manually on the basis of intravascular location and CRIg and F4/80 coexpression. Blood velocity was quantified using the automated tracking function in Volocity. A protocol was generated using the "find objects" and "track" tasks from the measurement modality, and labeled RBCs were tracked over a 1-min video. RBCs were assigned to their respective locations in sinusoids or venules and static events were excluded. Only RBCs that were trackable over 100 timepoints were included at an imaging speed of 10 frames per second. Three-dimensional (3D) reconstructions were generated using the "3D opacity" function in Volocity. Quantification of intravascular macrophages was performed in ImageJ using the image calculator function to count cells coexpressing F4/80 and CX3CR1 in the selected region of interest (ROI; venule) and expressed as percentage of surface area covered. Analysis of KC morphology, expression of KC markers, distribution of *S. aureus*, frequency of Ms4a3-labeling, frequency of CLEC4F nuclear signal and cytosolic signal, and frequency of PKH labeling and KC chimerism in parabiosis experiments was performed in Volocity software using the object counting tool. Three to five randomly selected FOV were assessed, the total number of KCs, and number of KCs expressing respective markers were quantified and expressed as percentage of total per FOV. The analysis was restricted to KC located in liver sinusoids. Volumetric analysis was done using Volocity software. Liver macrophage volume was assessed using computer-generated 3D renderings of 1-μm *z*-stack images. KC-like syncytia were identified on the basis of intravascular location in large vessels and coexpression of F4/80 and CRIg. Sinusoidal macrophage subsets were identified on the basis of their intravascular location in sinusoids and their respective expression of F4/80, CRIg, and TIM-4. We assessed five FOVs per mouse for KC-like syncytia and five randomly selected FOVs per mouse for macrophage subtypes in sinusoids. HSC and sinusoidal macrophage interaction was quantified using the "extended focus" function with *z*-stack images in Volocity software. A protocol was generated, using the "find objects" task from the measurement modality to identify all HSCs in the FOV. Second, an ROI was manually generated for each macrophage. Third, using the "intersect" task, objects and ROIs were combined to assess the surface area of individual macrophages covered by HSCs. The surface area of macrophages interacting with HSCs were then normalized to percentage of cell surface in contact with HSCs. Three to five FOVs were assessed per mouse. For all imaging quantifications, multiple FOV were assessed and averaged to count each mouse as $n = 1$. CRIg expression in human cirrhosis was quantified using ImageJ in automated fashion using the "analyze particles" task. Three to five randomly selected FOV were chosen, and coexpression of CRIg and IBA1 was then enumerated. Cell size in human liver tissues was analyzed as previously described (54). In brief, single-field images were acquired using a Zeiss Axio Observer 7. Large-field scanned images were stitched and further processed in ImageJ. Hyperstacks were concatenated and aligned. Cell identification, counting, distribution, and size measurement were performed using CellProfiler.

RNA sequencing

Eight samples in total (five controls and three CCl_4-treated) were analyzed for bulk RNA-seq. After isolation of hepatic leukocytes, 2×10^5 of embryonic KCs (F4/80^{hi}CD11b^{lo}Ms4a3^{-}) were purified by fluorescence-activated cell sorting and placed into 500 μl of QIAzol lysis buffer (QIAGEN, Venlo, Netherlands). Total RNA was isolated using a RNeasy Plus micro kit (QIAGEN, Venlo, Netherlands) and sent to the Centre for Health Genomics facility at the University of Calgary, where the RNA-seq was performed using a NextSeq sequencer (Illumina, San Diego, CA, USA). Total RNA was preprocessed to remove ribosomal RNA using NEBNext Ultra II Directional RNA Library Prep Kit for Illumina with rRNA depletion (New England Biolabs, Ipswich, MA, USA). The libraries were quantified by quantitative polymerase chain reaction using the KAPA Library Quantification Kit for Illumina with equimolar pooling and sequenced on a NextSeq 500 High output 75-cycle run. All samples passed the quality control for the library prep and sequence run metrics.

Analysis of RNA-seq data

Five control samples and three CCl_4-treated samples (including one pool of two unique treated samples owing to minimum sequencing library input requirements) were quantified using Kallisto v0.42.4 (55) against the NCBI RefSeq murine transcriptome (GRCm38). A linear regression model was used to determine differentially expressed transcripts using treatment (0 or 1) as the experimental variable. Differentially expressed genes were considered those with a false discovery rate (Benjamini-Hochberg–corrected $P < 0.05$) under the Wald test in Sleuth v0.30.0 (56) and a minimum of 40 raw reads mapped. Differentially expressed transcripts were subsequently annotated and analyzed using Ingenuity Pathway Analysis (Qiagen, Redwood City, CA) Core Analysis. DESeq2 was used to apply a negative binomial generalized linear model to determine differentially expressed genes between control and treatment groups (57). DESeq2 was used to generate PCA plots, and heatmaps were generated using the ComplexHeatmap package (58).

Analysis of single-cell RNA-seq data

The dataset used to interrogate gene expression of macrophages and monocytes in the CCl_4 model

was previously published (*12*) and is accessible in the GEO database (GSE136103). In this study, mice were first treated with CCl_4. Liver leukocytes were sorted and then processed for droplet-based single-cell RNA-seq. Mononuclear phagocytes were clustered according to the initial study in five clusters. Genes of interest were analyzed and displayed.

Statistical analysis

The sample size was predetermined before investigation on the basis of experience with similar hypothesis testing experiments. All data are represented as individual values with mean ± SD unless otherwise indicated in the figure legend. Generally, each data point is one biological replicate (*n*). The number of times an experiment was independently repeated (*N*) and the number of biological replicates (*n*) is stated in the figure legends. Statistical analyses were performed using GraphPad Prism (Version 8.0). Normality distribution was assessed and comparison of two groups was done either with unpaired Student's *t* test (two-tailed) or Mann-Whitney test, as appropriate. When assessing three or more groups, one-way analysis of variance (ANOVA) with Tukey's post hoc test was performed. Statistical significance was set at $P < 0.05$.

REFERENCES AND NOTES

1. C. N. Jenne, P. Kubes, Immune surveillance by the liver. *Nat. Immunol.* **14**, 996–1006 (2013). doi: 10.1038/ni.2691; pmid: 24048121
2. T. Huby, E. L. Gautier, Immune cell-mediated features of non-alcoholic steatohepatitis. *Nat. Rev. Immunol.* **22**, 429–443 (2022). doi: 10.1038/s41577-021-00639-3; pmid: 34741169
3. F. Heymann, F. Tacke, Immunology in the liver—From homeostasis to disease. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 88–110 (2016). doi: 10.1038/nrgastro.2015.200; pmid: 26758786
4. J. Bonnardel *et al*., Stellate cells, hepatocytes, and endothelial cells imprint the Kupffer cell identity on monocytes colonizing the liver macrophage niche. *Immunity* **51**, 638–654.e9 (2019). doi: 10.1016/j.immuni.2019.08.017; pmid: 31561945
5. Y. Iwakiri, V. Shah, D. C. Rockey, Vascular pathobiology in chronic liver disease and cirrhosis – current status and future directions. *J. Hepatol.* **61**, 912–924 (2014). doi: 10.1016/j.jhep.2014.05.047; pmid: 24911462
6. J. Gracia-Sancho, G. Marrone, A. Fernández-Iglesias, Hepatic microcirculation and mechanisms of portal hypertension. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 221–234 (2019). doi: 10.1038/s41575-018-0097-3; pmid: 30568278
7. D. Schuppan, N. H. Afdhal, Liver cirrhosis. *Lancet* **371**, 838–851 (2008). doi: 10.1016/S0140-6736(08)60383-9; pmid: 18328931
8. M. J. Lane, R. B. Jeffrey Jr., D. S. Katz, Spontaneous intrahepatic vascular shunts. *AJR Am. J. Roentgenol.* **174**, 125–131 (2000). doi: 10.2214/ajr.174.1.1740125; pmid: 10628467
9. E. M. Remer, G. A. Motta-Ramirez, J. M. Henderson, Imaging findings in incidental intrahepatic portal venous shunts. *AJR Am. J. Roentgenol.* **188**, W162–W167 (2007). doi: 10.2214/AJR.05.1115; pmid: 17242223
10. M. Praktiknjo *et al*., Total area of spontaneous portosystemic shunts independently predicts hepatic encephalopathy and mortality in liver cirrhosis. *J. Hepatol.* **72**, 1140–1150 (2020). doi: 10.1016/j.jhep.2019.12.021; pmid: 31954206
11. T. Su *et al*., Single-cell transcriptomics reveals zone-specific alterations of liver sinusoidal endothelial cells in cirrhosis. *Cell. Mol. Gastroenterol. Hepatol.* **11**, 1139–1161 (2021). doi: 10.1016/j.jcmgh.2020.12.007; pmid: 33340713
12. P. Ramachandran *et al*., Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* **575**, 512–518 (2019). doi: 10.1038/s41586-019-1631-3; pmid: 31597160
13. I. Mederacke *et al*., Fate tracing reveals hepatic stellate cells as dominant contributors to liver fibrosis independent of its aetiology. *Nat. Commun.* **4**, 2823 (2013). doi: 10.1038/ncomms3823; pmid: 24264436
14. B. G. Surewaard *et al*., Identification and treatment of the *Staphylococcus aureus* reservoir in vivo. *J. Exp. Med.* **213**, 1141–1151 (2016). doi: 10.1084/jem.20160334; pmid: 27325887
15. B. G. J. Surewaard, P. Kubes, Measurement of bacterial capture and phagosome maturation of Kupffer cells by intravital microscopy. *Methods* **128**, 12–19 (2017). doi: 10.1016/j.ymeth.2017.05.004; pmid: 28522327
16. M. Guilliams, C. L. Scott, Liver macrophages in health and disease. *Immunity* **55**, 1515–1529 (2022). doi: 10.1016/j.immuni.2022.08.002; pmid: 36103850
17. K. Y. Helmy *et al*., CRIg: A macrophage complement receptor required for phagocytosis of circulating pathogens. *Cell* **124**, 915–927 (2006). doi: 10.1016/j.cell.2005.12.039; pmid: 16530040
18. Z. Zeng *et al*., CRIg functions as a macrophage pattern recognition receptor to directly bind and capture blood-borne Gram-positive bacteria. *Cell Host Microbe* **20**, 99–106 (2016). doi: 10.1016/j.chom.2016.06.002; pmid: 27345697
19. A. Remmerie *et al*., Osteopontin expression identifies a subset of recruited macrophages distinct from Kupffer cells in the fatty liver. *Immunity* **53**, 641–657.e14 (2020). doi: 10.1016/j.immuni.2020.08.004; pmid: 32888418
20. J. S. Seidman *et al*., Niche-specific reprogramming of epigenetic landscapes drives myeloid cell diversity in nonalcoholic steatohepatitis. *Immunity* **52**, 1057–1074.e7 (2020). doi: 10.1016/j.immuni.2020.04.001; pmid: 32362324
21. S. Tran *et al*., Impaired Kupffer cell self-renewal alters the liver response to lipid overload during non-alcoholic steatohepatitis. *Immunity* **53**, 627–640.e5 (2020). doi: 10.1016/j.immuni.2020.06.003; pmid: 32562600
22. Z. Liu *et al*., Fate mapping via Ms4a3-expression history traces monocyte-derived cells. *Cell* **178**, 1509–1525.e19 (2019). doi: 10.1016/j.cell.2019.08.009; pmid: 31491389
23. E. Mass *et al*., Specification of tissue-resident macrophages during organogenesis. *Science* **353**, aaf4238 (2016). doi: 10.1126/science.aaf4238; pmid: 27492475
24. M. Sakai *et al*., Liver-derived signals sequentially reprogram myeloid enhancers to initiate and maintain Kupffer cell identity. *Immunity* **51**, 655–670.e8 (2019). doi: 10.1016/j.immuni.2019.09.002; pmid: 31587991
25. C. L. Scott *et al*., The transcription factor ZEB2 is required to maintain the tissue-specific identities of macrophages. *Immunity* **49**, 312–325.e5 (2018). doi: 10.1016/j.immuni.2018.07.004; pmid: 30076102
26. A. Guillot *et al*., Bile acid–activated macrophages promote biliary epithelial cell proliferation through integrin $\alpha v\beta 6$ upregulation following liver injury. *J. Clin. Invest.* **131**, e132305 (2021). doi: 10.1172/JCI132305; pmid: 33724957
27. K. R. Karlmark *et al*., Hepatic recruitment of the inflammatory $Gr1^+$ monocyte subset upon liver injury promotes hepatic fibrosis. *Hepatology* **50**, 261–274 (2009). doi: 10.1002/hep.22950; pmid: 19554540
28. S. Yona *et al*., Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity* **38**, 79–91 (2013). doi: 10.1016/j.immuni.2012.12.001; pmid: 23273845
29. F. F. Hoyer *et al*., Tissue-specific macrophage responses to remote injury impact the outcome of subsequent local immune challenge. *Immunity* **51**, 899–914.e7 (2019). doi: 10.1016/j.immuni.2019.10.010; pmid: 31732166
30. A. S. Neupane *et al*., Patrolling alveolar macrophages conceal bacteria from the immune system to maintain homeostasis. *Cell* **183**, 110–125.e11 (2020). doi: 10.1016/j.cell.2020.08.020; pmid: 32888431
31. R. Wiest, M. Lawson, M. Geuking, Pathological bacterial translocation in liver cirrhosis. *J. Hepatol.* **60**, 197–209 (2014). doi: 10.1016/j.jhep.2013.07.044; pmid: 23993913
32. L. Helming, S. Gordon, Molecular mediators of macrophage fusion. *Trends Cell Biol.* **19**, 514–522 (2009). doi: 10.1016/j.tcb.2009.07.005; pmid: 19733078
33. H. A. Crosby, P. F. Lalor, E. Ross, P. N. Newsome, D. H. Adams, Adhesion of human haematopoietic ($CD34^+$) stem cells to human liver compartments is integrin and CD44 dependent and modulated by CXCR3 and CXCR4. *J. Hepatol.* **51**, 734–749 (2009). doi: 10.1016/j.jhep.2009.06.021; pmid: 19703720
34. Y. Hou, Q. Zou, R. Ge, F. Shen, Y. Wang, The critical role of $CD133^+CD44^{+/high}$ tumor cells in hematogenous metastasis of liver cancers. *Cell Res.* **22**, 259–272 (2012). doi: 10.1038/cr.2011.139; pmid: 21862973
35. B. McDonald *et al*., Interaction of CD44 and hyaluronan is the dominant mechanism for neutrophil sequestration in inflamed liver sinusoids. *J. Exp. Med.* **205**, 915–927 (2008). doi: 10.1084/jem.20071765; pmid: 18362172
36. C. Shi *et al*., Monocyte trafficking to hepatic sites of bacterial infection is chemokine independent and directed by focal intercellular adhesion molecule-1 expression. *J. Immunol.* **184**, 6266–6274 (2010). doi: 10.4049/jimmunol.0904160; pmid: 20435926
37. O. Krenkel, F. Tacke, Liver macrophages in tissue homeostasis and disease. *Nat. Rev. Immunol.* **17**, 306–321 (2017). doi: 10.1038/nri.2017.11; pmid: 28317925
38. T. Langhans, Ueber Riesenzellen mit wandständigen Kernen in Tuberkeln und die fibröse Form des Tuberkels. *Virchows Arch.* **42**, 382–404 (1868). doi: 10.1007/BF02006420
39. M. T. Quinn, I. A. Schepetkin, Role of NADPH oxidase in formation and function of multinucleated giant cells. *J. Innate Immun.* **1**, 509–526 (2009). doi: 10.1159/000228158; pmid: 20375608
40. M. Guilliams, C. L. Scott, Does niche competition determine the origin of tissue-resident macrophages? *Nat. Rev. Immunol.* **17**, 451–460 (2017). doi: 10.1038/nri.2017.42; pmid: 28461703
41. A. M. Moon, A. G. Singal, E. B. Tapper, Contemporary epidemiology of chronic liver disease and cirrhosis. *Clin. Gastroenterol. Hepatol.* **18**, 2650–2666 (2020). doi: 10.1016/j.cgh.2019.07.060; pmid: 31401364
42. A. Martini, E. Ceranto, A. Gatta, P. Angeli, P. Pontisso, Occult liver disease burden: Analysis from a large general practitioners' database. *United European Gastroenterol. J.* **5**, 982–986 (2017). doi: 10.1177/2050640617696402; pmid: 29163964
43. J. Lewis, Histopathology of granulomatous liver disease. *Clin. Liver Dis. (Hoboken)* **11**, 77–80 (2018). doi: 10.1002/cld.695; pmid: 30992794
44. H. Zhu, H. C. Bodenheimer Jr., D. J. Clain, A. D. Min, N. D. Theise, Hepatic lipogranulomas in patients with chronic liver disease: Association with hepatitis C and fatty liver disease. *World J. Gastroenterol.* **16**, 5065–5069 (2010). doi: 10.3748/wjg.v16.i40.5065; pmid: 20976843
45. Y. Y. Pang *et al*., agr-Dependent interactions of *Staphylococcus aureus* USA300 with human polymorphonuclear neutrophils. *J. Innate Immun.* **2**, 546–559 (2010). doi: 10.1159/000319855; pmid: 20829608
46. B. G. Surewaard *et al*., Inactivation of staphylococcal phenol soluble modulins by serum lipoprotein particles. *PLOS Pathog.* **8**, e1002606 (2012). doi: 10.1371/journal.ppat.1002606; pmid: 22457627
47. B. Shrum *et al*., A robust scoring system to evaluate sepsis severity in an animal model. *BMC Res. Notes* **7**, 233 (2014). doi: 10.1186/1756-0500-7-233; pmid: 24725742
48. J. F. Deniset, B. G. Surewaard, W. Y. Lee, P. Kubes, Splenic $Ly6G^{high}$ mature and $Ly6G^{int}$ immature neutrophils contribute to eradication of *S. pneumoniae*. *J. Exp. Med.* **214**, 1333–1350 (2017). doi: 10.1084/jem.20161621; pmid: 28424248
49. M. Honda *et al*., Perivascular localization of macrophages in the intestinal mucosa is regulated by Nr4a1 and the microbiome. *Nat. Commun.* **11**, 1329 (2020). doi: 10.1038/s41467-020-15068-4; pmid: 32165624
50. C. Deppermann *et al*., Tacrolimus impairs Kupffer cell capacity to control bacteremia: Why transplant recipients are susceptible to infection. *Hepatology* **73**, 1967–1984 (2021). doi: 10.1002/hep.31499; pmid: 32761929
51. W. Y. Lee *et al*., An intravascular immune response to *Borrelia burgdorferi* involves Kupffer cells and iNKT cells. *Nat. Immunol.* **11**, 295–302 (2010). doi: 10.1038/ni.1855; pmid: 20228796
52. J. Zindel *et al*., Primordial GATA6 macrophages function as extravascular platelets in sterile injury. *Science* **371**, eabe0595 (2021). doi: 10.1126/science.abe0595; pmid: 33674464
53. J. Schindelin *et al*., Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). doi: 10.1038/nmeth.2019; pmid: 22743772
54. A. Guillot *et al*., Mapping the hepatic immune landscape identifies monocytic macrophages as key drivers of steatohepatitis and cholangiopathy progression. *Hepatology* **78**, 150–166 (2023). doi: 10.1097/HEP.0000000000000270; pmid: 36630995
55. N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016). doi: 10.1038/nbt.3519; pmid: 27043002
56. H. Pimentel, N. L. Bray, S. Puente, P. Melsted, L. Pachter, Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat. Methods* **14**, 687–690 (2017). doi: 10.1038/nmeth.4324; pmid: 28581496

57. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014). doi: 10.1186/s13059-014-0550-8; pmid: 25516281
58. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016). doi: 10.1093/bioinformatics/btw313; pmid: 27207943
59. M. Peiseler *et al*., Kupffer cell-like syncytia replenish resident macrophage function in fibrotic liver, Dryad (2023); https://doi.org/10.5061/dryad.3ffbg79pp.

ACKNOWLEDGMENTS

We thank L. Zbytnuik, M. Newton, and T. Nussbaumer for excellent technical support. We thank K. Poon for assistance with flow cytometry. We thank all current and past Kubes lab members for constant support and feedback. The authors thank L. Babes for help with IVM. We thank the Centre for Health Genomics and Informatics for help with RNA-seq. **Funding:** M.P. is supported by the German Research Foundation (DFG, PE2737/1-1) and is a member of the Berlin Institute of Health (BIH) Clinician Scientist Program; B.A.D. is supported by the Beverly Phillip's fellowship; J.Z. is supported by the Swiss National Science Foundation (SNSF P1BEP3_181164); J.D. and B.G.J.S. are supported by fellowships from the Canadian Institute of Health Research (CIHR); F.V.S.C. is supported by a fellowship from Canadian Institute of Health Research (MFE-176551); R.S. is a Cancer Research Institute Irvington Fellow supported by the Cancer Research Institute (CRI4653); Y.N. is a Robert Black Fellow of the Damon Runyon Cancer Research Foundation, DRG-2401-20; K.M. is funded by the Canadian Institute of Health Research (CIHR); and P.K. is supported by the NSERC Discover grant (RGPIN/07191-2019), the Heart and Stroke Foundation of Canada, CIHR, and Canada Research Chairs Program. **Author contributions:** M.P. designed and performed the experiments, analyzed data, wrote the manuscript, and designed figures. B.A.D. performed flow cytometry and imaging experiments, analyzed data, and designed figures. P.K. supervised the study, designed experiments, obtained funding, and wrote the manuscript. J.Z. analyzed RNA-seq data and helped with image data analysis. B.G.J.S. performed imaging experiments and helped design experiments. W.-Y.L. performed imaging experiments. Y.N. conducted flow cytometry experiments. F.V.S.C. performed lung imaging. R.S. performed imaging experiments. J.D. helped design experiments. C.O. and K.M. provided GF mice and assisted with experiments. P.G.M. performed electron microscopy experiments, and P.G.M. and M.Am. helped analyze electron microscopy data. F.H. analyzed flow experiments and acquired human images and helped with image analysis. C.P., S.M., A.G., A.B., A.N., R.T., and M.Al. provided human samples, performed immunofluorescence staining, and acquired microscopy images. F.T. provided human samples and gave critical input. Z.L. and F.G. provided critical materials and valuable input. J.A. performed bioinformatics and helped design figures. P.M.K.G. performed RNA-seq, bioinformatics analysis, and data curation. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Raw sequencing data have been deposited in the GEO database under GSE226946. Additional sequencing data used are available in the GEO database under GSE136103. The *Lrat*^{cre} mice were kindly provided by R. Schwabe (Columbia University, New York) under a material transfer agreement with the University of Calgary. All data underlying the results are deposited in Dryad (*59*). All other data are available in the main text or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq5202
Figs. S1 to S7
Table S1
Movies S1 to S8
MDAR Reproducibility Checklist

Submitted 13 April 2022; resubmitted 8 March 2023
Accepted 13 July 2023
10.1126/science.abq5202

RESEARCH ARTICLE SUMMARY

GREEN NUDGES

Reducing single-use cutlery with green nudges: Evidence from China's food-delivery industry

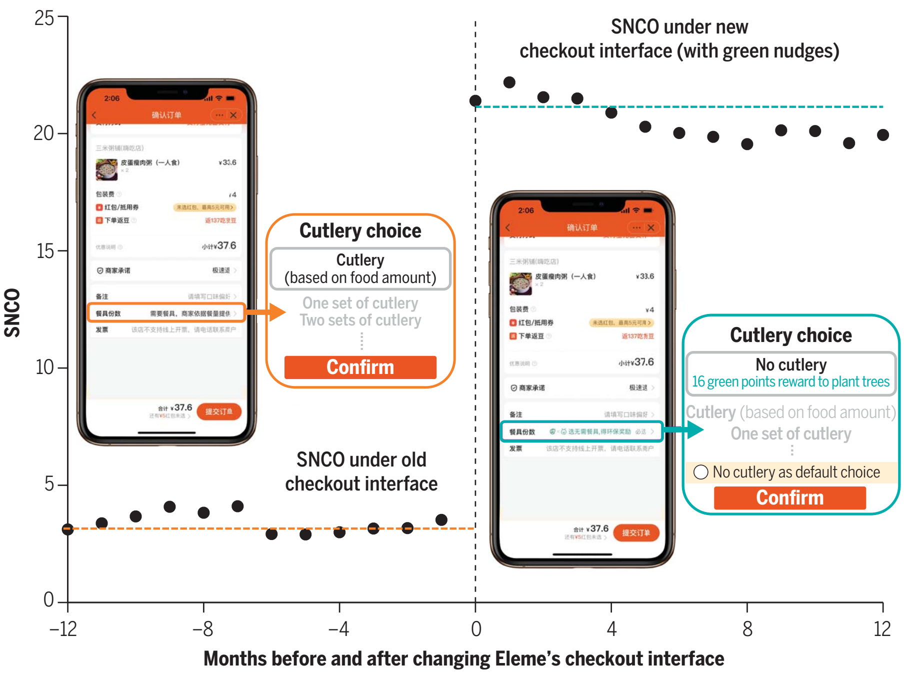
Guojun He*, Yuhang Pan, Albert Park, Yasuyuki Sawada, Elaine S. Tan

INTRODUCTION: Plastic waste is a global environmental threat that endangers marine and freshwater ecosystems worldwide. More recently, as food-delivery services became increasingly popular during the COVID-19 pandemic, the surge in plastic waste generated by single-use cutlery (SUC) has become a key environmental concern. Yet effective policies that control SUC waste are largely nonexistent, and it is important to find ways to encourage individuals to reduce their SUC consumption. Using data from China, the world's largest producer and consumer of SUC, we investigated how green nudges can affect individuals' cutlery decisions when placing food-delivery orders.

RATIONALE: From 2019 to 2020, three Chinese cities (Beijing, Shanghai, and Tianjin) introduced regulations that prohibited online food-delivery companies from including SUC unless it was explicitly requested. To comply with the regulations, Alibaba's food-delivery company, Eleme, changed its app in the following ways: (i) by adding a pop-up window that required customers to explicitly choose the number of SUC sets to be included with their orders, (ii) by setting the default for this pop-up window to be "no cutlery," and (iii) by providing a small nonpecuniary incentive––several Ant Forest green points––to those who chose the "no cutlery" option. The green points do not have a monetary value, but if one accumulates enough points (roughly by placing more than 1000 online food orders), they can be redeemed in exchange for planting a real tree (under the customer's name) in a desert area in China. The changes in the app's user interface embody the concept of "nudges" from behavioral economics and social psychology, which describes approaches that change the choice environment (or choice architecture) or provide indirective information to influence the behaviors and decision-making processes of individuals. Using customer-level data from Alibaba in 10 cities from 2019 to 2020, we compared behavioral differences between the customers in the "nudged" cities and those in the control cities before and after the introduction of green nudges.

RESULTS: The green nudges, on average, increased an individual's share of no-cutlery orders by 20.1 percentage points, which was a 648% increase relative to the baseline group. Meanwhile, the green nudges incentivized a large portion of individuals to somewhat change their behaviors rather than encouraging only a small portion to change their behaviors substantially. Women, older individuals, frequent food-delivery-service users, and wealthy individuals were more responsive to the green nudges. Importantly, Alibaba's business performance was not affected by the green nudges, suggesting that this could be a highly cost-effective way to reduce SUC waste. Additional mechanism analyses revealed that the default change and increased salience of the no-cutlery option were the main drivers of the observed behavioral changes, whereas the incentive to accumulate green points to plant trees played a relatively muted role. We estimate that if green nudges were applied to all of China, more than 21.75 billion sets of SUC could be saved annually, which is equivalent to preventing the generation of 3.26 million metric tons of plastic waste and saving 5.44 million trees.

CONCLUSION: Our study provides compelling evidence that nudges can be a powerful tool for changing behaviors. It also suggests that private sector and platform companies can provide highly cost-effective solutions to promote prosocial behaviors among their customers. In this study, the costs of implementing the green nudges were almost negligible (i.e., several hours of work to redesign the user interface), yet the aggregated environmental benefits were tremendous. We thus recommend that other online food-delivery platforms, such as DoorDash and Uber Eats, try similar green nudges to reduce global plastic waste. ■



Green nudges reduce SUC. The graph illustrates the trends in the share of no-cutlery orders (SNCO) among Alibaba's Eleme customers in three cities (Beijing, Shanghai, and Tianjin) before and after changing the app's checkout interface, with dashed orange and teal lines indicating the average SNCO. Compared with the old interface, the new interface has three nudging components: (i) a pop-up window that requires customers to explicitly choose the number of SUC sets to be included with their orders, (ii) the default for this pop-up window set to "no cutlery," and (iii) a small nonpecuniary incentive—some Ant Forest green points—that is given to those who choose the no-cutlery option.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: gjhe@hku.hk
Cite this article as G. He *et al.*, *Science* **381**, eadd9884 (2023). DOI: 10.1126/science.add9884

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.add9884

RESEARCH ARTICLE

GREEN NUDGES

Reducing single-use cutlery with green nudges: Evidence from China's food-delivery industry

Guojun He[1,2]*, Yuhang Pan[3], Albert Park[4,5,6,7], Yasuyuki Sawada[8,9], Elaine S. Tan[4]

Rising consumer demand for online food delivery has increased the consumption of disposable cutlery, leading to plastic pollution worldwide. In this work, we investigate the impact of green nudges on single-use cutlery consumption in China. In collaboration with Alibaba's food-delivery platform, Eleme (which is similar to Uber Eats and DoorDash), we analyzed detailed customer-level data and found that the green nudges—changing the default to "no cutlery" and rewarding consumers with "green points"—increased the share of no-cutlery orders by 648%. The environmental benefits are sizable: If green nudges were applied to all of China, more than 21.75 billion sets of single-use cutlery could be saved annually, equivalent to preventing the generation of 3.26 million metric tons of plastic waste and saving 5.44 million trees.

Plastic waste is a global threat that endangers marine and freshwater ecosystems worldwide (*1*–*3*). In 2021, more than 400 million metric tons of plastic waste were produced worldwide, and it is predicted that the world's plastic waste growth will continue to outpace the efforts to reduce plastic pollution in the coming decades (*3*). More recently, as food-delivery services became increasingly popular during the COVID-19 pandemic, the surge in plastic waste generated by single-use cutlery (SUC) has become a key environmental challenge faced by many countries (*4*).

Reducing SUC waste in the food-delivery industry is particularly important in China, the world's largest producer and consumer of SUC. As of 2019, more than 540 million Chinese were active users of food-delivery services and each day consumed more than 50 million sets of SUC that were not adequately recycled or disposed of (*5*). SUC in China typically includes a plastic fork, a plastic spoon, a pair of wooden chopsticks, and a napkin. Consequently, SUC usage in China not only generates large amounts of plastic waste (i.e., used forks and spoons) but also consumes a large number of trees (discarded chopsticks and napkins), which reduces forest area, threatens ecosystems, and ultimately poses substantial health risks to humans (*6*).

To reduce SUC consumption, policy-makers in China set a target of reducing SUC usage in food deliveries by 30% by 2025. Online platforms were required to establish plans to achieve that target (*7*). From 2019 to 2020, Shanghai, Beijing, and Tianjin further introduced pilot city-wide regulations that prohibited online food-delivery companies from including SUC in their orders unless it was explicitly requested by customers. To comply with the new rules, online food-delivery platforms redesigned their app user interfaces and required that customers explicitly request SUC when placing an order.

In this study, we collaborated with Alibaba's online food-ordering platform Eleme to evaluate the effectiveness of their efforts to reduce SUC consumption. Eleme, which is similar to Uber Eats and DoorDash, is China's second-largest food-delivery company, with more than 753 million users in 2022. Hereafter, we refer to the platform simply as "Alibaba." To comply with city regulations, Alibaba changed its app in the following ways: (i) by adding a pop-up window that required customers to explicitly choose the number of SUC sets with their orders, (ii) by setting the default for this pop-up window to be "no cutlery," and (iii) by providing a small nonpecuniary incentive––some "Ant Forest green points" (hereafter "green points")—to those choosing the no-cutlery option (see supplementary note 1). The green points do not have a monetary value, but if one accumulates enough points (roughly, by placing more than 1000 online food orders), they can be redeemed in exchange for planting a real tree (under the customer's name) in a desert area in China.

The changes in the Alibaba app's user interface embody the concept of "nudging" from behavioral economics and social psychology, which describes approaches that change the choice environment (or choice architecture) or provide indirective information to influence the behaviors and decision-making of individuals (*8*, *9*). The key feature of nudges is that they do not limit the choices available to individuals, nor do they change the monetary incentives (*8*, *9*). In the past two decades, nudges have been used in many social domains (*10*), including green nudges that promote pro-environmental behaviors (*11*, *12*). However, to what extent green nudges can be effective remains controversial (*11*, *13*–*16*), and there are few opportunities for researchers to examine the impacts of green nudges at scale and for an extended period of time.

Using detailed customer-level data from Alibaba (*17*), we investigated how Alibaba's green nudges affected individuals' cutlery decisions. Specifically, we applied a difference-in-differences model to the data and compared the food-ordering behaviors of individuals in the pilot (hereafter, "treated") cities with green nudges with those of individuals in the "control" cities without green nudges before and after Alibaba changed the app interface (Materials and methods). Our research findings not only provide compelling evidence for how green nudges can affect pro-environmental behaviors but also generate important implications for the understanding of nudges in general.

Specifically, this study has several distinctive features that distinguish it from the vast literature on nudges. First, there is an ongoing debate about why nudges work and under what conditions they work (*18*–*21*), and the rich data in this study allowed us to explore the underlying mechanisms through which the green nudges affect individuals' behaviors. Second, with a few exceptions (*22*, *23*), most studies in the nudge literature focus on the short-term impacts; by contrast, our study can follow individuals' repeated decisions, explore the persistence of the nudging effect, and examine whether nudges work in situations where individuals can easily switch to other options and set other options as the default. Examining the medium- and long-term effects is particularly important in our setting because short-term analysis might exaggerate the treatment effects. Third, nudges have been criticized for being manipulative because they are not always transparent and can take advantage of individuals' ignorance or lack of awareness (*24*–*29*). However, the green nudges we studied here are easy to understand and completely transparent to users (i.e., the pop-up window that requires individuals to make explicit choices). Fourth, the panel data of individuals allow us to estimate individual-specific effects, which can inform us of how many people are "nudgeable" and the entire distribution of the nudge effects. Finally, few policies target plastic-waste production at

[1]Faculty of Business and Economics, University of Hong Kong, Hong Kong SAR, China. [2]Institute for Climate and Carbon Neutrality, University of Hong Kong, Hong Kong SAR, China. [3]Institute for Global Health and Development, Peking University, Beijing, China. [4]Asian Development Bank, Metro Manila, Philippines. [5]Department of Economics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, China. [6]Division of Social Science, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, China. [7]Division of Public Policy, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, China. [8]Faculty of Economics, University of Tokyo, Tokyo, Japan. [9]Asian Development Bank Institute, Tokyo, Japan.
***Corresponding author. Email: gjhe@hku.hk**

the consumer level, except charges on plastic bags (30–33). Thus, our results can help policymakers better understand whether it is feasible to reduce consumer plastic consumption by regulating platform companies.

Methods

Changes in Alibaba's app user interface

Alibaba redesigned its app's user interface to comply with the new environmental regulations in Shanghai (from 1 July 2019), Beijing (from 1 May 2020), and Tianjin (from 1 December 2020). Figure 1 depicts the changes. Fig. 1A shows the old user interface: The options for SUC were placed at the end of the food-ordering page, with the default being "with cutlery (number of sets based on the amount of food)." Restaurants typically provided SUCs freely and sometimes even gave more SUCs than needed to avoid negative reviews (34).

In the new user interface, as shown in Fig. 1B, when a customer in a treated city placed an order and clicked the check-out button, a window emerged that required the customer to explicitly choose the number of sets of SUC. Importantly, on this pop-up window, the default became "no cutlery," and customers had to scroll down the screen to choose a different option. Furthermore, to prevent the window from appearing in their future orders, a customer could set any of the options as their new device-level default by clicking the "set as default" button. The no-cutlery option came with a small nonpecuniary incentive, that is, the green points. When a customer placed an order without SUC, 16 green points would be awarded to the person, which could be stored and accumulated in Alibaba's payment app, Alipay. When consumers accumulated 16,000 green points, they could spend them to request that Alibaba plant a real tree named after the consumer in a desert region of China. There are also other ways to obtain green points under Alibaba's platform, such as walking more, taking more public transportation, selling used items, and so on. The green points have helped plant more than 223 million trees in China from 2016 to 2020 (35). Getting a tree planted only by placing no-cutlery orders is challenging. According to our data, the average number of monthly no-cutlery orders for a consumer was only 0.77 from 2019 to 2020.

We could identify three nudging components of the changes to the user interface: (i) a pop-up window that increased the salience of available cutlery options to the customers, (ii) the change of the default selection to "no cutlery," and (iii) the inclusion of a small nonpecuniary incentive (green points) as part of the new default option. The effect we estimated in this study thus came from a mixture of different nudging incentives rather than just one. According to the nudge taxonomy suggested by Münscher *et al.* (36), nudges can be classified into three types: those that provide decision information (like reframing, increasing visibility, and providing social reference points), those that provide decision assistance (like reminders and commitment devices), and those that change the decision structure (like changing the defaults, the option-related efforts, the range of composition of options, and the option consequences). In our setting, the salience intervention provided decision information, that is, it made the cutlery options more explicit, transparent, and visible. The other two nudging elements altered the decision structure, including a standard change in the default (i.e., "no cutlery") and a change in the option consequences (i.e., green points for planting trees). The green-points incentive could also be viewed as a symbolic award, which provided individuals with an intrinsic value that went beyond and exceeds that of an expendable or disposable item (37).

Our data included each user's monthly food-ordering history from 1 January 2019 to 31 December 2020 in 10 major Chinese cities (17). These include the three treated cities with green nudges (i.e., Beijing, Shanghai, and Tianjin) and the seven control cities without the nudges (Qingdao, Xi'an, Guangzhou, Nanjing, Hangzhou, Wuhan, and Chengdu). Among these cities, we randomly sampled about 200,000 "active" users (i.e., those who placed at least one order between 2019 and 2020). We then

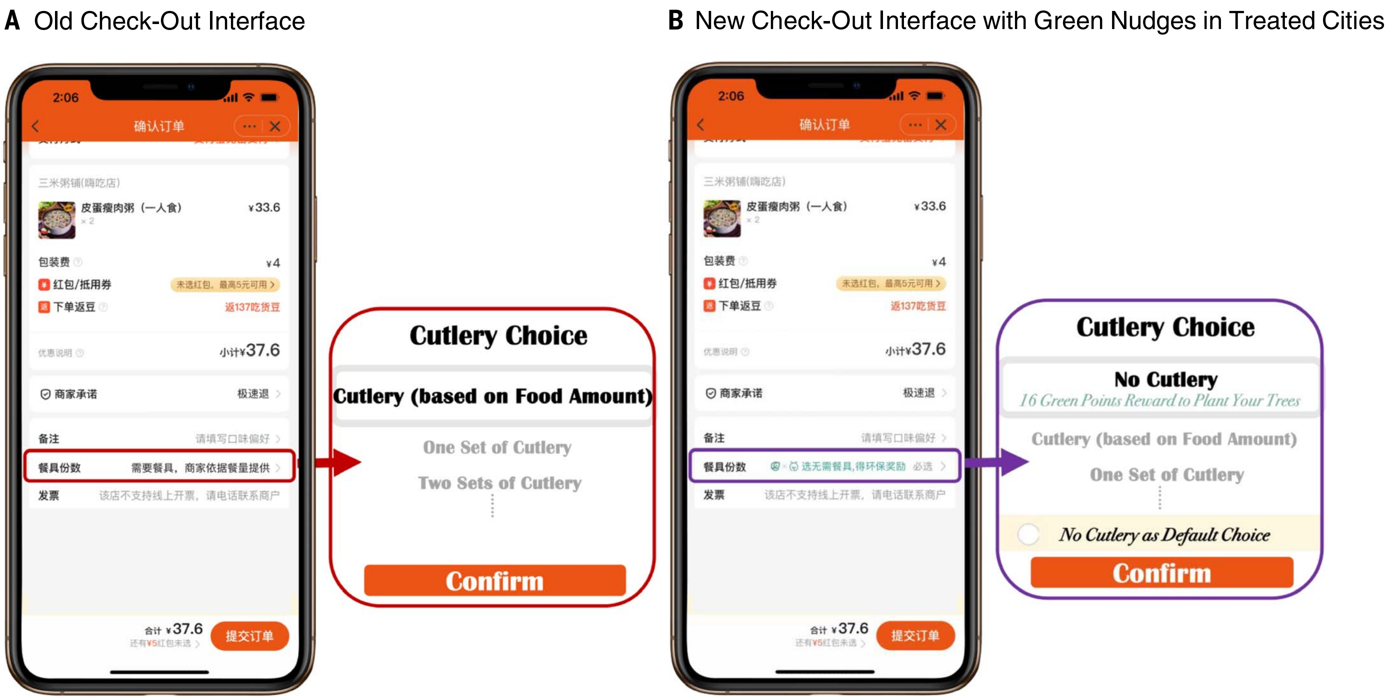


Fig. 1. Green-nudge interventions at the Eleme application interface. (**A**) The old interface, in which the default cutlery option at the check-out page was preset as "with cutlery." Users had to click on the pop-up window and scroll to the bottom to choose the no-cutlery option. (**B**) The new interface in the treated cities. A window about cutlery automatically popped up during checkout and required the consumer to choose the number of sets of SUC. The default cutlery option was "no cutlery." Users could make any option their default by clicking the "set as default" button (so that this window would not pop up in their future orders). The no-cutlery option came with a small nonpecuniary incentive, that is, 16 Ant Forest green points. Contents are translated by the authors.

compared the behavioral differences between the users in the treated cities and the users in the control cities before and after the introduction of green nudges using a difference-in-differences model (Eq. 1 in Materials and methods). We asked the following four research questions (RQs):

RQ 1. Did the green nudges affect individuals' no-cutlery orders and did the effects persist? We hypothesized that the green nudges should increase the no-cutlery orders because both the default change and the green points incentivized them to do so.

RQ 2. Did the green nudges negatively affect the platform's business performance? Because the new interface did not reduce the options available to customers and offered additional green-point incentives, we hypothesized that the user experience of the platform would not deteriorate, and, therefore, its business performance would not be negatively affected.

RQ 3. Which types of customers, if any, were more likely to respond to the green nudges? The answer depended on the costs and benefits of choosing the no-cutlery option for different individuals, as well as their general "nudgeability." For example, those who were more environmentally conscious and ate at home could be more nudgeable than others.

RQ 4. Were the aggregate environmental benefits large? The answer to this question depended on the extent to which the green nudges affected SUC consumption and how much waste reduction this implies.

Results

Green nudges substantially increased the no-cutlery orders

Figure 2A depicts the changes in the raw data. We observed that the share of no-cutlery orders significantly increased in the treated cities after the intervention, whereas the share remained relatively unchanged throughout the study period in the control cities.

We then estimated the difference-in-differences model (Eq. 1 in Materials and methods) and summarized the results in Fig. 2B. We observed that the green nudges increased the share of no-cutlery orders by 19.3 percentage points in Shanghai, 21.2 percentage points in Beijing, and 20.4 percentage points in Tianjin (regression results in table S1). On average, the green nudges increased an individual's share of no-cutlery orders by 20.1 percentage points. Because the share of no-cutlery orders before the treatment was around 3.1 percentage points, the frequency of no-cutlery orders increased by 648% (20.1 divided by 3.1 and multiplied by 100). The impacts of green nudges observed in this study are significantly larger than those found in the literature, as summarized in table S2.

We conducted an event-study analysis to examine the plausibility of the parallel trend assumption in the difference-in-differences model (Eq. 2 in Materials and methods). We observed that before the introduction of the green nudges, the control and treated cities followed essentially the same trends (Fig. 2C and table S3). Immediately after the introduction of green nudges, however, the share of no-cutlery orders in the treated cities significantly increased. These results suggest that

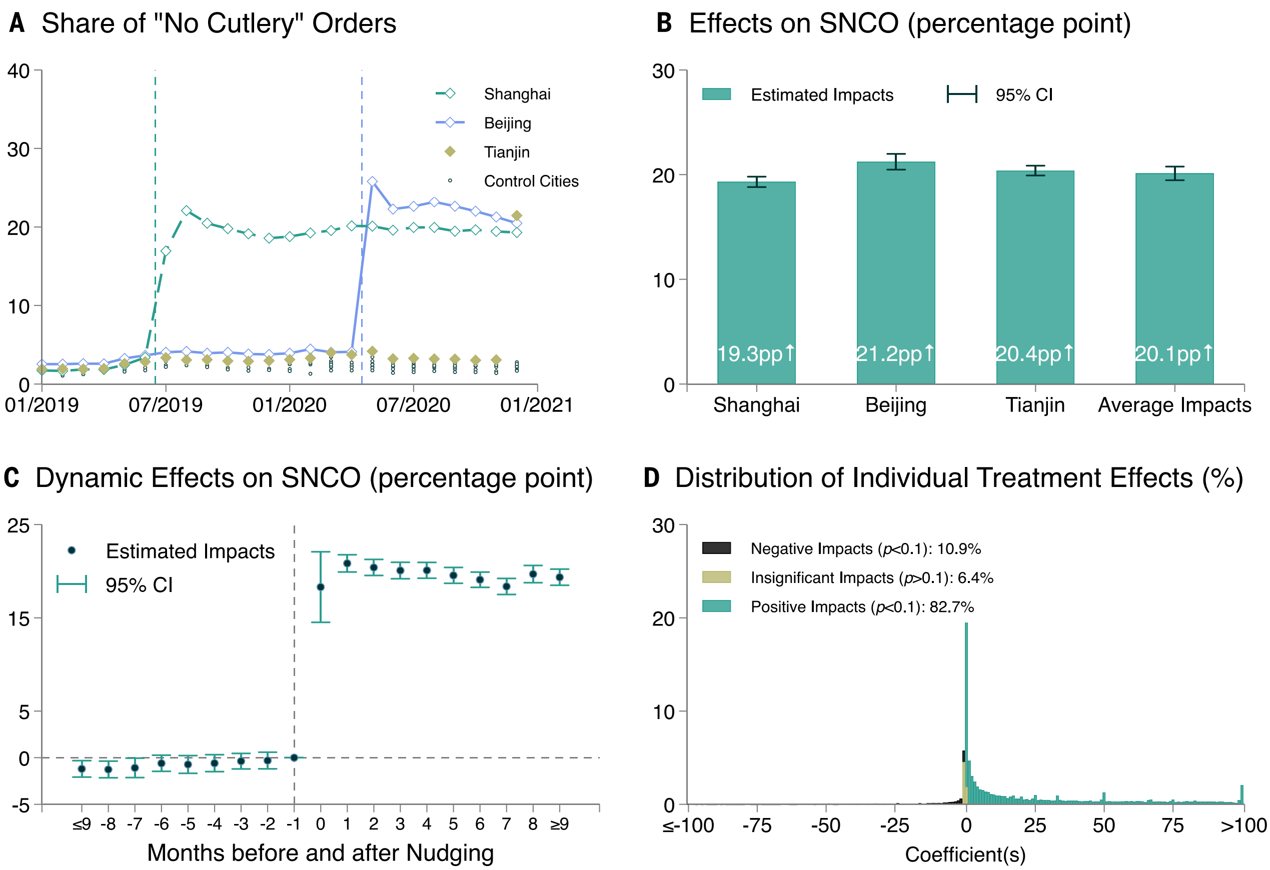


Fig. 2. Investigation of the impacts of green nudges. (**A**) Trends in the share of no-cutlery orders (SNCO). The vertical green and blue dashed lines indicate introductions of the green nudges in Shanghai and Beijing, respectively. The green nudges were implemented in December 2020 in Tianjin. White circles refer to SNCO in the other seven cities in the control group. (**B**) The estimated impacts of green nudges on SNCO using individual-level data. The first three bars present the estimated effects of green nudges in each treated city separately. The fourth bar reports the average treatment effect across all cities. (**C**) The event-study results, where the reference group in the regression is 1 month before the introduction of the green nudges. CI, confidence interval. (**D**) The distribution of individual treatment-effect estimates for all the consumers in the treated cities. The *x* and *y* axes indicate the effect size and density of the estimates, respectively.

in the absence of green nudges, customers in the two groups of cities would follow similar SUC consumption patterns. In supplementary note 2, we further show that the media and public attention to plastic waste and pollution did not change significantly when the city-wide regulations were introduced, which ensures that our findings are driven by the app changes, rather than by the regulations directly.

We further estimated the individual treatment effects and summarized the estimates in Fig. 2D (Eq. 1 in Materials and methods). We made several observations. First, nearly 83% of the individuals in the treated cities responded positively to the green nudges (positive and statistically significant at the 10% level). Second, cutlery choices did not change much for about 6% of the individuals in the treated cities (close to zero and statistically nonsignificant at the 10% level). Third, the remaining 11% of the customers were nudge defiers and behaved in the opposite direction to what was encouraged. Figure S1 further reveals that most nudge defiers were customers who had previously placed no-cutlery orders, suggesting that a small share of environmentally conscious people actually disliked being nudged. Fourth, although most of the population responded positively to the green nudges, the density of the positive estimates was negatively correlated with the effect magnitude. Finally, most of the consumers did not set "no cutlery" as their default option (otherwise, most estimates would be close to or equal to one) nor did most of them set "with cutlery" as their default option (otherwise most estimates would be close to or equal to zero).

Green nudges did not harm business performance

One concern regarding the green nudges was that some customers might dislike the new choice architecture and became less likely to use the Alibaba platform for food deliveries. If this were to happen, such nudges could become unsustainable because the platform would lose revenue in the long run, with less environmentally friendly rivals gaining a competitive advantage.

Figure 3 depicts the changes in customers' total spending on the platform each month and the total number of orders they placed. Both the total order amount and the total number of orders followed the same trend in the treated and control cities (results in table S4). The total spending and total number of orders substantially dropped in early 2020 in all the cities, which was caused by the Chinese government's stringent anticontagion policies during the COVID-19 pandemic. Therefore, we conclude that the impact of green nudges on the platform's business performance is negligible and statistically nonsignificant.

Certain subpopulations responded more to the green nudges

We examined the heterogeneous impacts of nudges for different subpopulations. The results are summarized in Fig. 4 (regression results in tables S5 to S11).

Several patterns emerged. First, the effects were larger for women. Specifically, the no-cutlery orders of women increased by 21.4 percentage points after the nudges were introduced compared with an 18.4 percentage point increase for men. Second, the effects also differed across age groups. For customers between the ages of 18 and 24, green nudges only increased the share of no-cutlery orders by 11.9 percentage points, whereas for the middle-aged and elderly, the increase was as much as 30 to 34 percentage points. Each difference was statistically significant. Third, frequent users (i.e., those that rarely cooked) responded less to the green nudges than infrequent users. Fourth, there were significant differences across wealth groups, as defined by the value of individuals' cell phones. For those using high-value cell phones (>CNY 8000 or USD 1151 as of January 2020), green nudges increased the share of no-cutlery orders by 22.2 percentage points, which was 3.92 percentage points higher than for those using low-value cell phones (≤CNY 1900 or USD 273). Fifth, and relatedly, those who ordered more-expensive meals (based on per-order expenditure) responded more to the green nudges. Thus, whether measured by cell phone value or meal price, more-affluent individuals were more responsive to the green nudges. Finally, the green nudges increased the share of no-cutlery orders by 24 percentage points for individuals who had previously placed no-cutlery orders before the intervention, which was 4 to 5 percentage points higher than for those who had never placed a no-cutlery order before. It is possible that this difference stems from a difference in inherent environmental consciousness.

Aggregate environmental benefits were nontrivial

Based on the findings reported in the previous sections, we estimated that the green nudges, in total, reduced the number of SUCs in the

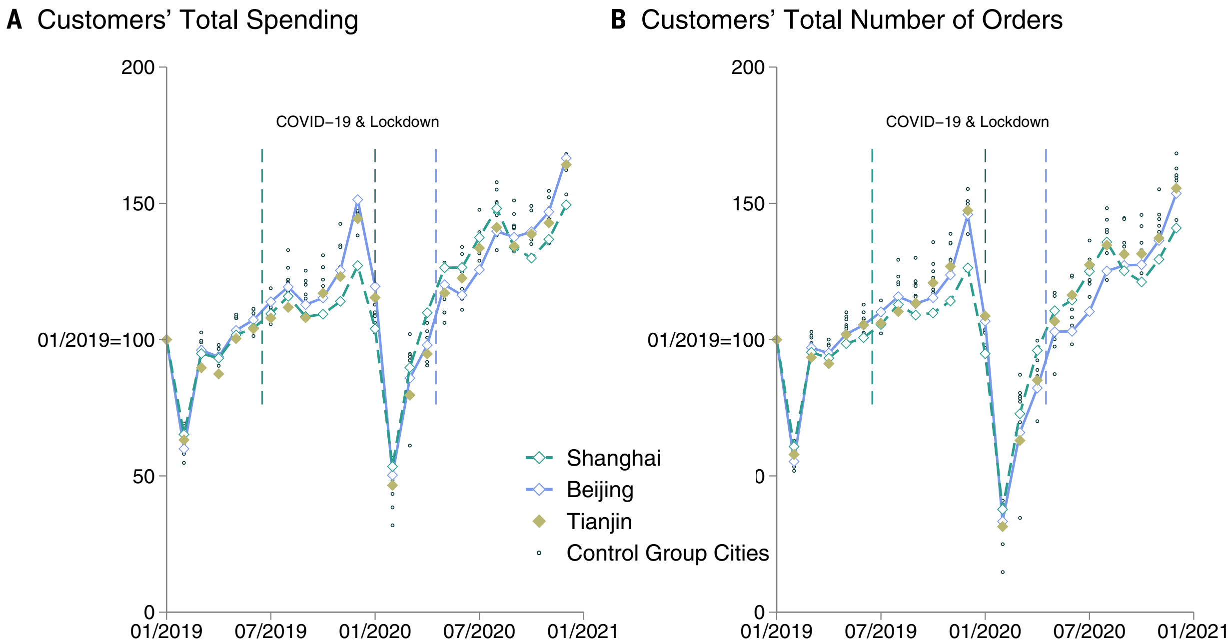


Fig. 3. Changes in food-delivery business. Changes in (**A**) customers' total spending and (**B**) customers' total number of orders. The vertical green and blue dashed lines indicate the introductions of the green nudges in Shanghai and Beijing, respectively. The green nudges were implemented in December 2020 in Tianjin.

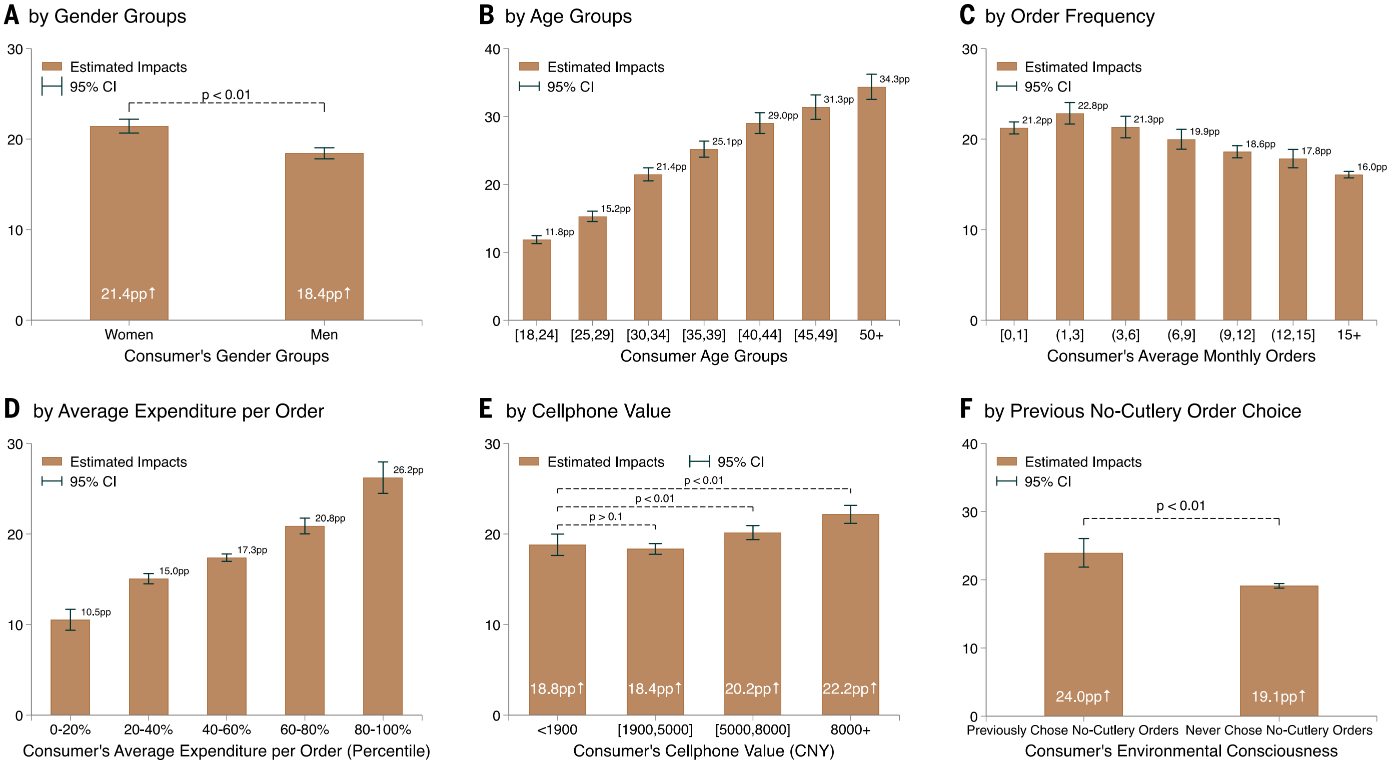


Fig. 4. Heterogeneous impacts of green nudges on different subpopulations. Subsample analysis based on consumer characteristics (**A**) gender, (**B**) age, (**C**) order frequency, (**D**) average expenditure per order, (**E**) cellphone value, and (**F**) previous no-cutlery orders before the introduction of green nudges. Each bar represents a separate regression. Values on the *y* axes are percentage points (pp). Corresponding results are reported in tables S5 to S10.

treated cities by more than 225.33 million sets during the 27 months of our study (18 months in Shanghai, 8 months in Beijing, and 1 month in Tianjin) (Fig. 5). Based on the weight and composition of a typical set of SUC, this reduction in SUC consumption may have prevented the generation of 4506.52 metric tons of waste and saved 56,333 trees (supplementary note 3). Further, if all the green points rewarded through the green nudges were used to plant trees, 112,665 additional trees could have been planted by Alibaba. If Alibaba introduced the green nudges to the entire country, more than 8.7 billion sets of SUC would be saved annually in China (based on 2020 data; supplementary note 3). Additionally, if all food-delivery services in the country adopted green nudges, the total SUC consumption would decrease by 21.75 billion sets (Fig. 5), which is equivalent to eliminating 3.26 million metric tons of plastic waste and saving 5.44 million trees.

These numbers should be interpreted as the upper-bound environmental benefits for two reasons. First, in reality, despite customers choosing the no-cutlery option when placing orders, some restaurants nevertheless provided SUC. This would occur, for example, when the restaurant was too busy to customize its orders or was concerned that the consumers might have mistakenly chosen the no-cutlery option. Second, people often wasted their green points or did not accumulate enough points to plant trees (supplementary note 1).

Mechanisms behind the behavioral changes

Alibaba's app changes included three nudging components: default change, green-points incentive, and increased salience of the SUC options. We conducted additional analyses to understand the contributions of different nudging components to the observed nudge effects.

First, our analyses described in supplementary note 4 show that the incentive to accumulate green points was unlikely to be the main driving force for individuals' behavioral changes. This is because (i) the effect of the green nudges did not depend on an individual's pretreatment green points (table S12), (ii) the effect was similar for those who had previously planted a tree and for those who had never done so (table S13), (iii) the effect was similar for users whose green points were closer to the threshold of being able to plant a tree (table S14), and (iv) the green nudges did not significantly increase individuals' total green points and most customers were not actively accumulating green points by forgoing SUC (table S15).

We then examined the underlying channels through which the default change could change individuals' SUC choices. The previous literature suggested several theories that might explain the default effect, including switching cost, present-biased preferences and procrastination (*38*–*40*), reference dependence (*41*–*44*), inattention to default change (*29*, *45*–*49*), and endorsement effect (*42*, *50*, *51*). We carefully examined these theories in supplementary note 5 and reached the following conclusions. First, because the switching cost was negligible in our setting, it was unlikely to be the driving force of the default effect. Second, the consumers' decisions did not involve complicated monetary or discounting calculations and they needed to make decisions repeatedly, so present-biased preferences and procrastination also fell short in explaining the observed behavioral changes. Third, reference dependence and loss aversion could not explain the default effect in our setting either. Under the no-cutlery default, consumers were not choosing to forgo something of value; therefore, loss aversion vis-à-vis the default choice did not apply. Fourth, given the repeated nature of online food orders and the new default popping up for each order, inattention to the default change also failed to explain the observed effects.

Therefore, we conclude that much of the observed effects should be driven by the endorsement effect brought by the default change and increased salience. Seeing the new user interface, individuals inferred that the default

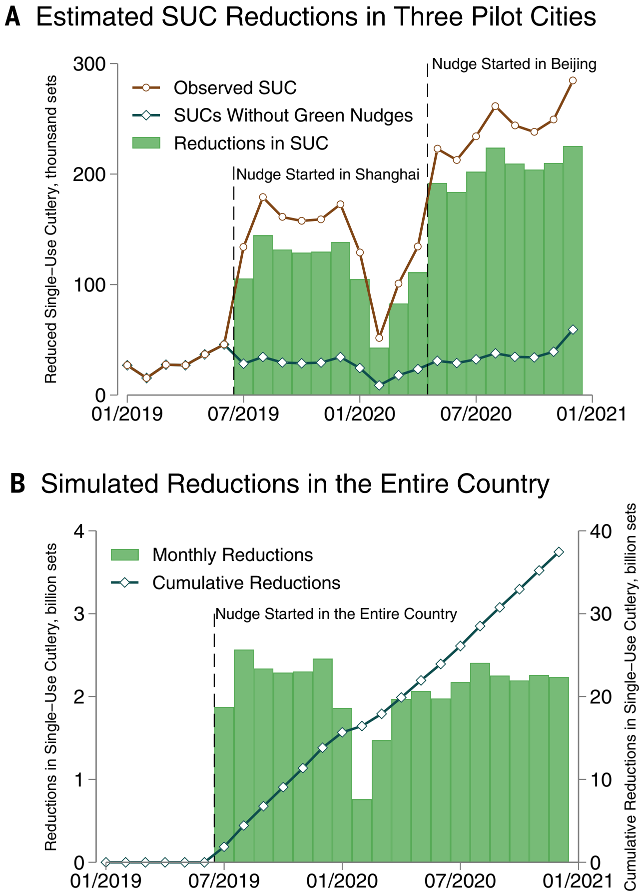


Fig. 5. Estimated reduction in SUC consumption from green nudges. (**A**) Estimated SUC reductions in pilot cities. The green bars illustrate the counterfactual reductions in SUC in Beijing and Shanghai after the introduction of green nudges. (**B**) Estimated potential reduction in SUC for all of China if the green nudges had been introduced nationally in July 2019.

option was encouraged and endorsed by the platform and thus became more likely to choose it. Meanwhile, the pop-up window made the decision about SUC more salient and reminded consumers that the no-cutlery option was available, so the increased salience of the new default could also have played a role.

Discussion

Using data from China, we show in this work that green nudges can substantially reduce plastic waste generated by the food-delivery industry. Changing the default to "no cutlery" and rewarding consumers with green points on Alibaba's food-delivery platform increased the share of no-cutlery orders by 648%, and the effect persisted. If green nudges were applied to all of China, more than 21.75 billion sets of SUC would be saved annually—equivalent to 20.4% of plastic waste reduction in the food delivery industry—which would eliminate 3.26 million metric tons of plastic waste and save 5.44 million trees.

There are several implications of this study. First, it offers compelling evidence that nudges can be a powerful tool to change behaviors. This finding contrasts with some recent reviews that show that nudges are largely ineffective, especially after accounting for publication bias (*52*, *53*). Two factors may help explain the large differences between our estimates and those in the literature. First, Alibaba's green nudges were introduced into a new domain where most individuals had not been affected by other policies or other types of encouragement, so the baseline take-up rate was low. Second, the endorsement effect might be particularly strong in our setting because adhering to the default does not incur a notable cost for most individuals. We thus conclude that the effectiveness of nudges is highly context-specific, and more nudges should be explored in new domains.

The results also help us better understand how individuals adjusted their behaviors. For example, the impacts of nudging were greatest in the first few months after the introduction and decreased slightly over time, indicating that some customers later decided to override the default no-cutlery option. This is consistent with previous literature on the decaying effect of default (*23*). We also found that women, middle-aged or older individuals, frequent users, affluent consumers, and environmentally conscious customers were more responsive to the green nudges. This information can be used to improve the targeting of future pro-environmental interventions. Further, the individual treatment effects revealed that the green nudges incentivized a large portion of individuals to somewhat change their behaviors rather than encouraging only a small portion to change their behaviors substantially and that people seemed to get prompted every time and occasionally opted not to have SUCs, rather than setting their own defaults. We also documented a somewhat puzzling phenomenon: A small portion of individuals were nudge defiers and responded in the opposite direction to what was encouraged. The reasons for such antinudging behaviors remain largely unknown.

Another important finding is that green nudges did not negatively affect Alibaba's business, suggesting that they can be a highly cost-effective tool to promote individuals' pro-environmental behaviors. For Alibaba, the cost of implementing the green nudges was trivial because it only required several software engineers to redesign the user interface. Yet the environmental benefits were tremendous. We thus recommend that other online food-delivery platforms, such as DoorDash and Uber Eats, try similar green nudges to reduce global plastic waste. More generally, we think that the private sector and platform companies can play a powerful role in promoting prosocial behaviors among their customers. Better alignment between their corporate social responsibilities and ecofriendly initiatives could bring about far-reaching impacts to our planet.

We conclude by noting the caveats of this study and future directions. First, because of the lack of data from other food-delivery platforms, we cannot examine whether Alibaba's green nudges had positive spillovers for transactions on their platforms. If positive spillovers exist, the benefits of green nudges would be even greater. Second, through field research, we learned that some restaurants provided cutlery to customers even when the no-cutlery option was chosen. This could undermine the power of the green nudges and reduce the environmental benefits associated with customers' pro-environmental behaviors. Future research is warranted to find ways to encourage restaurants to better comply with customers' requests. Finally, although the aggregated environmental benefits of the green nudges were nontrivial, a substantial portion of the solid and plastic waste in the food-delivery industry comes from packaging (i.e., using disposable food containers and plastic bags for delivery services). In typical Chinese food-delivery orders, packaging waste accounts for more than 80% of the food-delivery waste (*54*). Additional policies should be introduced to better control the waste generated in the packaging process.

Materials and methods

Data

We obtained data from Eleme, Alibaba's food-ordering and -delivery platform, which is similar

to DoorDash and Uber Eats. From June 2019 to June 2022, the total number of monthly active users on Eleme increased from 709 million to 753 million, making it the second-largest platform in China (next to Meituan) (55).

The data included 197,062 randomly selected users' monthly food-ordering history, their green-points history, their tree-planting records, and their personal characteristics from 1 January 2019 to 31 December 2021 in 10 major Chinese cities. All of the consumers in the sample set placed at least one food delivery order during the research period.

Among the 10 cities, there were three treated cities, including Beijing, Shanghai, and Tianjin. Seven cities were the control cities, including Qingdao, Xi'an, Guangzhou, Nanjing, Hangzhou, Wuhan, and Chengdu. According to China's provincial statistical yearbook, these cities had a combined total population of 157.36 million as of 2020.

Eleme provided monthly data on each customer's food-ordering history, including the number of orders, the number of no-cutlery orders, and total expenditures. Based on this information, we calculated the share of no-cutlery orders (SNCO) in percentage terms, that is, we divided the number of no-cutlery orders by the total number of orders and multiplied by 100. Data on green points included total rewarded green points, total rewarded green points from no-cutlery orders, and total harvested and/or accumulated green points. More details about the green points are discussed in supplementary note 1. Data on tree-planting records included each customer's cumulative and current-month tree-planting activities. Customers' personal characteristics included gender, age group, and approximate value of their cellphone, which was computed by Eleme's algorithm. The correlations between pretreatment SNCO behaviors and consumers' characteristics are summarized in table S16.

Model

We used a difference-in-differences approach to quantify the impact of the green nudges on individuals' SUC-ordering behaviors. Specifically, the following model was estimated:

$$SNCO_{ict} = \alpha + \beta^{*}Nudge_{ict} + \pi_i + \rho_t + \varepsilon_{ict} \quad (1)$$

where the dependent variable is the SNCO for individual i in city c in year-month t, which is calculated by $SNCO_{ict} = \frac{\text{Number of orders without SUC}}{\text{Total number of orders}} * 100$. If a customer did not make any food-delivery order in a specific month, we dropped this observation (rather the entire consumer's record) from the regression. $Nudge_{ict}$ is the treatment variable that equals one if individual i in city c received the green nudges in a specific year and month and is zero otherwise. We also controlled for individual fixed effects, π_i, and time (year-by-month) fixed effects, ρ_t. α is the intercept, β is the coefficient of interest, and ε_{ict} is the error term. Standard errors were two-way clustered at the individual level and at the city-year-month level (56).

The difference-in-differences model estimates the treatment effects of nudges, β, by comparing the changes in the share of no-cutlery orders in the treated cities (Beijing, Shanghai, and Tianjin) with changes in the control cities (seven cities in the control group) before and after the nudges were implemented. The identifying assumption was that, in the absence of the green nudges, customers in the treated cities would display similar SUC usage trends as customers in the control cities. We tested this assumption by examining the trends in SNCO for both treated and control cities before and after the introduction of the green nudges. Specifically, we estimated an event study specification as follows:

$$SNCO_{ict} = \alpha + \sum_{k\geq-9,k\neq-1}^{k=9} \beta_k * D_Nudge_{ict,k} + \pi_i + \rho_t + \varepsilon_{ict} \quad (2)$$

where $D_Nudge_{ict,k}$ is a set of dummy variables indicating the treatment status in different periods: For individual i in city c who was never nudged, $D_Nudge_{ict,k}$ was always set equal to zero; but for individual i in city c who received the green nudges, we first defined S_{ct} as the period during which the nudges were introduced and then we defined $D_Nudge_{ict,k} = 1$ if $t - S_{ct} = k$ and equal to zero otherwise, where $k \in [-9, 9]$. The dummy for $k = -1$ was omitted in the regression as the reference group. Therefore, β_k dynamically captured the impacts of green nudges from when they were first introduced until 9 months later. By including leads of treatment timing, we could test whether pretrends differed for treatment and control cities by examining whether green nudges had any impact on outcomes up to 9 months before the actual implementation.

To understand how each individual was affected by the green nudges relative to the control group, we estimated the individual treatment effects using Eq. 1 and only a subsample for the estimation: individuals in the control cities plus one individual in the treated city. The standard errors were then clustered at the individual level. We then repeated the regression for all the treated individuals, obtaining a large number of individual-specific treatment-effect estimates. Note that the individual effect estimates should be interpreted with caution because the statistical power was limited given the small sample size for each individual.

To investigate whether the treatment effect of the green nudges was stronger for users whose points were closer to the threshold of being able to plant a tree (and thus revealed a strong motivation to respond to the green-points incentive), we followed Grembi *et al.* (57) and Giambona and Ribas (58) and applied a difference-in-discontinuities specification using the following model:

$$SNCO_{ict} = \beta_1 Below_i + \beta_2 Diff_i + \beta_3 Below_i \cdot Diff_i + \gamma_0 Nudge_{ict} + \gamma_1 Below_i \cdot Nudge_{ict} + \gamma_2 Diff_i \cdot Nudge_{ict} + \gamma_3 Below_i \cdot Diff_i \cdot Nudge_{ict} + \varepsilon_{ict}$$

$$\text{subject to } -h \leq Diff_i \leq h \quad (3)$$

where $SNCO_{ict}$ is the share of no-cutlery orders for individual i in city c in year-month t. Dummy variable $Below_i$ was equal to one when individual i's points were below 16,000 points—because the minimum requirement for planting a tree was 16,000 points—before the treatment and equal to zero if individual i's points were above or equal to the threshold. $Diff_i$ is the running variable calculated by the difference between individual i's points and the threshold of 16,000 points before the treatment. A negative value for $Diff_i$ suggests that individual i's points are below 16,000. $Nudge_{ict}$ is the treatment variable that equals one if individual i in city c received the green nudges in a specific year and month and that equals zero otherwise. h is the bandwidth chosen by researchers to estimate the discontinuity. Individual fixed effects and year-month fixed effects were included in the regression. Standard errors were two-way clustered both at the individual level and at the city-year-month level. The difference-in-discontinuity estimator is γ_1, which measures the treatment effect on individuals who were below the threshold.

REFERENCES AND NOTES

1. J. R. Jambeck *et al*., Plastic waste inputs from land into the ocean. *Science* **347**, 768–771 (2015). doi: 10.1126/science.1260352; pmid: 25678662
2. L. C. M. Lebreton *et al*., River plastic emissions to the world's oceans. *Nat. Commun.* **8**, 15611 (2017). doi: 10.1038/ncomms15611; pmid: 28589961
3. S. B. Borrelle *et al*., Predicted growth in plastic waste exceeds efforts to mitigate plastic pollution. *Science* **369**, 1515–1518 (2020). doi: 10.1126/science.aba3656; pmid: 32943526
4. T. M. Adyel, Accumulation of plastic waste during COVID-19. *Science* **369**, 1314–1315 (2020). doi: 10.1126/science.abd9925; pmid: 32913095
5. Xinhua Net; http://www.xinhuanet.com/fortune/2019-12/19/c_1125362989.htm.
6. R. Qi, D. L. Jones, Z. Li, Q. Liu, C. Yan, Behavior of microplastics and plastic film residues in the soil environment: A critical review. *Sci. Total Environ.* **703**, 134722 (2020). doi: 10.1016/j.scitotenv.2019.134722; pmid: 31767311
7. National Development and Reform Commission, Ministry of Ecology and Environment, "Opinions on further strengthening the cleanup of plastic pollution" (2020); https://www.ndrc.gov.cn/xxgk/zcfb/tz/202001/t20200119_1219275.html.
8. R. H. Thaler, From cashews to nudges: The evolution of behavioral economics. *Am. Econ. Rev.* **108**, 1265–1287 (2018). doi: 10.1257/aer.108.6.1265
9. R. H. Thaler, Nudge, not sludge. *Science* **361**, 431–431 (2018). doi: 10.1126/science.aau9241; pmid: 30072515
10. R. H. Thaler, C. R. Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Penguin Books, 2009).
11. F. Carlsson, C. Gravert, O. Johansson-Stenman, V. Kurz, The use of green nudges as an environmental policy

instrument. *Rev. Environ. Econ. Policy* **15**, 216–237 (2021). doi: 10.1086/715524
12. C. Schubert, Green nudges: Do they work? Are they ethical? *Ecol. Econ.* **132**, 329–342 (2017). doi: 10.1016/j.ecolecon.2016.11.009
13. J. Wensing, V. Caputo, L. Carraresi, S. Bröring, The effects of green nudges on consumer valuation of bio-based plastic packaging. *Ecol. Econ.* **178**, 106783 (2020). doi: 10.1016/j.ecolecon.2020.106783
14. R. Croson, N. Treich, Behavioral environmental economics: Promises and challenges. *Environ. Resour. Econ.* **58**, 335–351 (2014). doi: 10.1007/s10640-014-9783-y
15. M. Akbulut-Yuksel, C. Boulatoff, The effects of a green nudge on municipal solid waste: Evidence from a clear bag policy. *J. Environ. Econ. Manage.* **106**, 102404 (2021). doi: 10.1016/j.jeem.2020.102404
16. D. Miłaszewicz, Survey results on using nudges for choice of green-energy supplier. *Energies* **15**, 2679 (2022). doi: 10.3390/en15072679
17. G. He, Y. Pan, A. Park, Y. Sawada, E. Tan, Data and code info for: Reducing single-use cutlery with green nudges: Evidence from China's food delivery industry. Dryad (2023); https://doi.org/10.5061/dryad.wm37pvmsx.
18. M. Fowlie *et al.*, Default effects and follow-on behaviour: Evidence from an electricity pricing program. *Rev. Econ. Stud.* **88**, 2886–2934 (2021). doi: 10.1093/restud/rdab018
19. J. Blumenstock, M. Callen, T. Ghani, Why do defaults affect behavior? Experimental evidence from Afghanistan. *Am. Econ. Rev.* **108**, 2868–2901 (2018). doi: 10.1257/aer.20171676
20. B. D. Bernheim, A. Fradkin, I. Popov, The welfare economics of default options in 401(k) plans. *Am. Econ. Rev.* **105**, 2798–2837 (2015). doi: 10.1257/aer.20130907
21. R. Chetty, J. N. Friedman, S. Leth-Petersen, T. H. Nielsen, T. Olsen, Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark. *Q. J. Econ.* **129**, 1141–1219 (2014). doi: 10.1093/qje/qju013
22. U. Liebe, J. Gewinner, A. Diekmann, Large and persistent effects of green energy defaults in the household and business sectors. *Nat. Hum. Behav.* **5**, 576–585 (2021). doi: 10.1038/s41562-021-01070-3; pmid: 33707659
23. A. Brandon, P. Ferraro, J. List, R. Metcalfe, M. Price, F. Rundhammer, "Do the effects of nudges persist? Theory and evidence from 38 natural field experiments," Working Paper w23277, National Bureau of Economic Research, Cambridge, MA, March 2017.
24. L. Bovens, "The ethics of nudge" in *Preference Change*, T. Grüne-Yanoff, S. O. Hansson, Eds. (Springer, 2009), pp. 207–219.
25. N. Dhingra, Z. Gorn, A. Kener, J. Dana, The default pull: An experimental demonstration of subtle default effects on preferences. *Judgm. Decis. Mak.* **7**, 69–76 (2012). doi: 10.1017/S1930297500001844
26. P. G. Hansen, A. M. Jespersen, Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy. *Eur. J. Risk Regul.* **4**, 3–28 (2013). doi: 10.1017/S1867299X00002762
27. H. Bruns, E. Kantorowicz-Reznichenko, K. Klement, M. Luistro Jonsson, B. Rahali, Can nudges be transparent and yet effective? *J. Econ. Psychol.* **65**, 41–59 (2018). doi: 10.1016/j.joep.2018.02.002
28. F. Ebeling, S. Lotz, Domestic uptake of green energy promoted by opt-out tariffs. *Nat. Clim. Chang.* **5**, 868–871 (2015). doi: 10.1038/nclimate2681
29. N. C. Smith, D. G. Goldstein, E. J. Johnson, Choice without awareness: Ethical and policy implications of defaults. *J. Public Policy Mark.* **32**, 159–172 (2013). doi: 10.1509/jppm.10.114
30. G. O. Thomas, E. Sautkina, W. Poortinga, E. Wolstenholme, L. Whitmarsh, The English plastic bag charge changed behavior and increased support for other charges to reduce plastic waste. *Front. Psychol.* **10**, 266 (2019). doi: 10.3389/fpsyg.2019.00266; pmid: 30863332
31. H. He, Effects of environmental policy on consumption: Lessons from the Chinese plastic bag regulation. *Environ. Dev. Econ.* **17**, 407–431 (2012). doi: 10.1017/S1355770X1200006X
32. S. Larcom, L. A. Panzone, T. Swanson, Follow the leader? Testing for the internalization of law. *J. Legal Stud.* **48**, 217–244 (2019). doi: 10.1086/699817
33. J. M. Cabrera, M. Caffera, A. Cid, Modest and incomplete incentives may work: Pricing plastic bags in Uruguay. *J. Environ. Econ. Manage.* **110**, 102525 (2021). doi: 10.1016/j.jeem.2021.102525
34. C. Lai, "Chongqing green group sues three delivery services over disposable cutlery," *Hong Kong Free Press*, 6 September 2017; https://hongkongfp.com/2017/09/06/chongqing-green-group-sues-three-delivery-services-disposable-cutlery/.
35. http://dse.rcees.cas.cn/dtxx/dtxxwz/202103/t20210331_632275.html
36. R. Münscher, M. Vetter, T. Scheuerle, A review and taxonomy of choice architecture techniques. *J. Behav. Decis. Making* **29**, 511–524 (2016). doi: 10.1002/bdm.1897
37. J. Gallus, Fostering public good contributions with symbolic awards: A large-scale natural field experiment at Wikipedia. *Manage. Sci.* **63**, 3999–4015 (2017). doi: 10.1287/mnsc.2016.2540
38. T. O'Donoghue, M. Rabin, Choice and procrastination. *Q. J. Econ.* **116**, 121–160 (2001). doi: 10.1162/003355301556365
39. J. J. Choi, D. Laibson, B. C. Madrian, A. Metrick, Optimal defaults. *Am. Econ. Rev.* **93**, 180–185 (2003). doi: 10.1257/000282803321947010
40. G. D. Carroll, J. J. Choi, D. Laibson, B. C. Madrian, A. Metrick, Optimal defaults and active decisions. *Q. J. Econ.* **124**, 1639–1674 (2009). doi: 10.1162/qjec.2009.124.4.1639; pmid: 20041043
41. R. H. Thaler, S. Benartzi, Save More Tomorrow™: Using behavioral economics to increase employee saving. *J. Polit. Econ.* **112**, S164–S187 (2004). doi: 10.1086/380085
42. B. C. Madrian, D. F. Shea, The power of suggestion: Inertia in 401(k) participation and savings behavior. *Q. J. Econ.* **116**, 1149–1187 (2001). doi: 10.1162/003355301753265543
43. K. M. Marzilli Ericson, A. Fuster, The endowment effect. *Annu. Rev. Econ.* **6**, 555–579 (2014). doi: 10.1146/annurev-economics-080213-041320
44. R. H. Thaler, Toward a positive theory of consumer choice. *J. Econ. Behav. Organ.* **1**, 39–60 (1980). doi: 10.1016/0167-2681(80)90051-7
45. R. Chetty, A. Looney, K. Kroft, Salience and taxation: Theory and evidence. *Am. Econ. Rev.* **99**, 1145–1177 (2009). doi: 10.1257/aer.99.4.1145
46. H. Allcott, Social norms and energy conservation. *J. Public Econ.* **95**, 1082–1095 (2011). doi: 10.1016/j.jpubeco.2011.03.003
47. D. Hirshleifer, S. S. Lim, S. H. Teoh, Driven to distraction: Extraneous events and underreaction to earnings news. *J. Finance* **64**, 2289–2325 (2009). doi: 10.1111/j.1540-6261.2009.01501.x
48. B. Bollinger, P. Leslie, A. Sorensen, Calorie posting in chain restaurants. *Am. Econ. J. Econ. Policy* **3**, 91–128 (2011). doi: 10.1257/pol.3.1.91
49. J. R. Kling, S. Mullainathan, E. Shafir, L. C. Vermeulen, M. V. Wrobel, Comparison friction: Experimental evidence from Medicare drug plans. *Q. J. Econ.* **127**, 199–235 (2012). doi: 10.1093/qje/qjr055; pmid: 22454838
50. C. R. M. McKenzie, M. J. Liersch, S. R. Finkelstein, Recommendations implicit in policy defaults. *Psychol. Sci.* **17**, 414–420 (2006). doi: 10.1111/j.1467-9280.2006.01721.x; pmid: 16683929
51. J. R. Brown, A. M. Farrell, S. J. Weisbenner, Decision-making approaches and the propensity to default: Evidence and implications. *J. Financ. Econ.* **121**, 477–495 (2016). doi: 10.1016/j.jfineco.2016.05.010
52. S. Mertens, M. Herberz, U. J. J. Hahnel, T. Brosch, The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2107346118 (2022). doi: 10.1073/pnas.2107346118; pmid: 34983836
53. M. Maier *et al.*, No evidence for nudging after adjusting for publication bias. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2200300119 (2022). doi: 10.1073/pnas.2200300119; pmid: 35858389
54. Y. Zhang, Z. Wen, Mapping the environmental impacts and policy effectiveness of takeaway food industry in China. *Sci. Total Environ.* **808**, 152023 (2022). doi: 10.1016/j.scitotenv.2021.152023; pmid: 34861304
55. Data sources are as follows: https://www.chyxx.com/industry/202009/896228.html#:~:text=2015%2D2019%E5%B9%B4%E9%A5%BF%E4%BA%86,%E5%A2%9E%E5%8A%A0%E4%BA%860.08%E4%BA%BF%E4%BA%BA%E3%80%82 and https://www.statista.com/statistics/1218045/china-leading-food-delivery-apps-based-on-monthly-active-users/#statisticContainer.
56. A. C. Cameron, J. B. Gelbach, D. L. Miller, Robust inference with multiway clustering. *J. Bus. Econ. Stat.* **29**, 238–249 (2011). doi: 10.1198/jbes.2010.07136
57. V. Grembi, T. Nannicini, U. Troiano, Do fiscal rules matter? *Am. Econ. J. Appl. Econ.* **8**, 1–30 (2016). doi: 10.1257/app.20150076
58. E. Giambona, R. P. Ribas, Unveiling the Price of Obscenity: Evidence from Closing Prostitution Windows in Amsterdam. *J. Policy Anal. Manage.* **42**, 677–705 (2023). doi: 10.1002/pam.22459

ACKNOWLEDGMENTS

We benefited from L. Goette's insights and thank seminar participants at the ADB project workshop for their comments and suggestions. We also thank two Alibaba project coordinators, R. Zhang and Y. Wang, for providing data. S. Kimura, D. Guo, A. Tan, J. Li, and Y. Quan provided excellent coordination and research assistance. **Funding:** We acknowledge funding support from the Research Grant Council of Hong Kong (theme-based research grant T31-603/21-N) and Peking University (early career grant 7101303264). The views expressed in this paper are those of the authors only and do not necessarily reflect the views and policies of the Asian Development Bank or its board of governors or the governments they represent. **Author contributions:** All authors contributed equally to this work and are listed alphabetically in the author list. Conceptualization: G.H., A.P., Y.S., E.S.T.; Methodology and data collection: Y.P., E.S.T.; Analysis: G.H., Y.P., A.P., Y.S., E.S.T.; Writing – original draft: G.H., Y.P.; Writing – review and editing: G.H., Y.P., A.P., Y.S., E.S.T. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The data usage guidelines and codes necessary to re-produce and extend all the tables from this research can be accessed through Dryad (*17*). The customer-level data contain private information and are thus stored by a designated server provided by Alibaba. Those who want to access the data for replication should file an application following the instructions provided by (*17*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add9884
Supplementary Text
Figs. S1 to S4
Tables S1 to S16
References (*59*–*71*)
MDAR Reproducibility Checklist

Submitted 18 July 2022; accepted 26 July 2023
10.1126/science.add9884

RESEARCH ARTICLE

COMMUNITY ECOLOGY

Widespread shifts in body size within populations and assemblages

Inês S. Martins[1,2]*, Franziska Schrodt[3], Shane A. Blowes[4,5], Amanda E. Bates[6], Anne D. Bjorkman[7,8], Viviana Brambilla[1,9], Juan Carvajal-Quintero[4,10], Cher F. Y. Chow[1], Gergana N. Daskalova[11], Kyle Edwards[12], Nico Eisenhauer[4,10], Richard Field[3], Ada Fontrodona-Eslava[1], Jonathan J. Henn[13,14], Roel van Klink[4,5], Joshua S. Madin[15], Anne E. Magurran[1], Michael McWilliam[1], Faye Moyes[1], Brittany Pugh[3,16], Alban Sagouis[4,5], Isaac Trindade-Santos[1,17], Brian J. McGill[18], Jonathan M. Chase[4,5], Maria Dornelas[1,2,9]

Biotic responses to global change include directional shifts in organismal traits. Body size, an integrative trait that determines demographic rates and ecosystem functions, is thought to be shrinking in the Anthropocene. Here, we assessed the prevalence of body size change in six taxon groups across 5025 assemblage time series spanning 1960 to 2020. Using the Price equation to partition this change into within-species body size versus compositional changes, we detected prevailing decreases in body size through time driven primarily by fish, with more variable patterns in other taxa. We found that change in assemblage composition contributes more to body size changes than within-species trends, but both components show substantial variation in magnitude and direction. The biomass of assemblages remains quite stable as decreases in body size trade off with increases in abundance.

The loss or gain of large organisms can have severe consequences for ecosystem functions in terms of total system biomass and metabolism, and thus food web energy fluxes (*1*). Anthropogenic changes to the biosphere are miniaturizing many communities (*1*–*3*) because of the selective removal of the largest individuals [population change, e.g., (*4*)] and the extinction of larger species [compositional change, e.g., (*5*)]. Population shifts in body size have been attributed to anthropogenic selection forces, including preferential exploitation of larger individuals (*6*, *7*) and climate change and habitat conversion (*8*–*10*). At the assemblage level, species compositional change is a major component of biodiversity change (*11*, *12*). Large-bodied species have been particularly susceptible to extinction after temperature shifts (in either direction) and human exploitation, largely because of their life history traits and lower abundances (*13*). However, shrinking trends of community body sizes are by no means universal, and when examined across many communities, some results suggest no net change in body size (*14*). Indeed, as species shift their ranges, some communities are gaining larger species, such as in arctic and high-elevation plant assemblages (*6*). Therefore, the prevalence of population and compositional body size changes and their implications for assemblage abundance and biomass are unknown.

Previous assessments have investigated either compositional components of body size change (*14*) or within-species population changes (*3*) or, when both components are examined, have focused on a single taxon (*15*). However, interactions of change in the two components can lead to nonintuitive outcomes in body size distributions through time. For example, the loss of competitors arising from compositional change can allow the remaining populations to increase in body size (*16*). Several scenarios of body size change can arise by simultaneously considering within-species population and compositional body size changes (Fig. 1). The two components can operate in the same direction (both increasing or decreasing body size) or change can involve only one component (change only in population or in composition but not both). Finally, change in one component can cancel out change in the other. Here, we explicitly set out to test to what extent each of the two components most commonly drives the changes in body size that we observed across taxa and regions and how often change in the two components occurs in the opposite directions.

To examine these scenarios, we used time series that recorded organism abundance and body size (biomass) data in the field and quantified the contribution of both components, compositional and within-species body size changes, to change in body size distributions across taxa. Specifically, we collated 5025 assemblages over 60 years (*17*), a time period of intensification of anthropogenic selection forces on the biosphere (*18*). These time series ranged from 5 to 56 years of surveys and covered 4292 species within six taxonomic groups (1971 fish species, 1201 plants species, 628 invertebrate species, 66 mammals species, 33 herpetofauna species, and 393 marine benthic organisms) from communities distributed across multiple regions of the world (fig. S1). The time series cover a variety of body sizes and body size change trends. We found body size shrinking in more assemblages than increasing across all datasets, as well as among time series that had stronger evidence for trends of change [lower *P* values; see Fig. 2, fig. S3, and (*19*) for details].

We quantified the prevalence of different components of change (Fig. 1) by decomposing body size change of all assemblages into compositional and within-species changes using an extension of the Price equation (*19*–*21*). The Price equation is a mathematical description of the relationship between statistical descriptors (mean and covariance) of selection and trait change (*22*). Although developed in an evolutionary context, this equation has direct application to the question of body size change through time if we equate competition and environmental filtering as selection [(*23*) but see also (*24*)]. By examining the type of covariance in these two components of change, we can determine the relative contributions of compositional and within-species change to the observed overall change (Fig. 1B).

Results and Discussion

Partitioning patterns of body size change

Our analysis shows that over all assemblages, body size is predominantly shrinking, with substantial variation in the balance of within-species change and compositional change (Fig. 3). Two-thirds of the 5025 assemblages decreased

[1]Centre for Biological Diversity, School of Biology, University of St Andrews, St Andrews KY16 9TH, Scotland. [2]Leverhulme Centre for Anthropocene Biodiversity, University of York, York YO10 5DD, UK. [3]School of Geography, University of Nottingham, University Park, Nottingham NG7 2RD. [4]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig 04103, Germany. [5]Department of Computer Science, Martin Luther University Halle-Wittenberg, Halle (Saale) 06099, Germany. [6]Department of Biology, University of Victoria, Victoria, British Columbia, Canada. [7]Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg 40530, Sweden. [8]Gothenburg Global Biodiversity Centre, Gothenburg 41319, Sweden. [9]MARE, Guia Marine Laboratory, Faculty of Sciences, University of Lisbon, Cascais 2750-374, Portugal. [10]Institute of Biology, Leipzig University, Leipzig 04103, Germany. [11]International Institute for Applied Systems Analysis (IIASA), Laxenburg 2361, Austria. [12]Department of Oceanography, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA. [13]Department of Evolution, Ecology, and Organismal Biology, University of California Riverside, Riverside, CA 92521, USA. [14]Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, CO 80303, USA. [15]Hawai'i Institute of Marine Biology, University of Hawai'i at Manoa, Kāne'ohe, Hawai'i 96744, USA. [16]University College London, School of Geography, Gower Street, London WC1E 6AE, UK. [17]Macroevolution Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1, Tancha, Onna-son, Kunigami-gun 904-0495, Okinawa, Japan. [18]School of Biology and Ecology and Mitchell Center for Sustainability Solutions, University of Maine, Orono, ME 04469, USA.
*Corresponding author. Email: istmartins@gmail.com

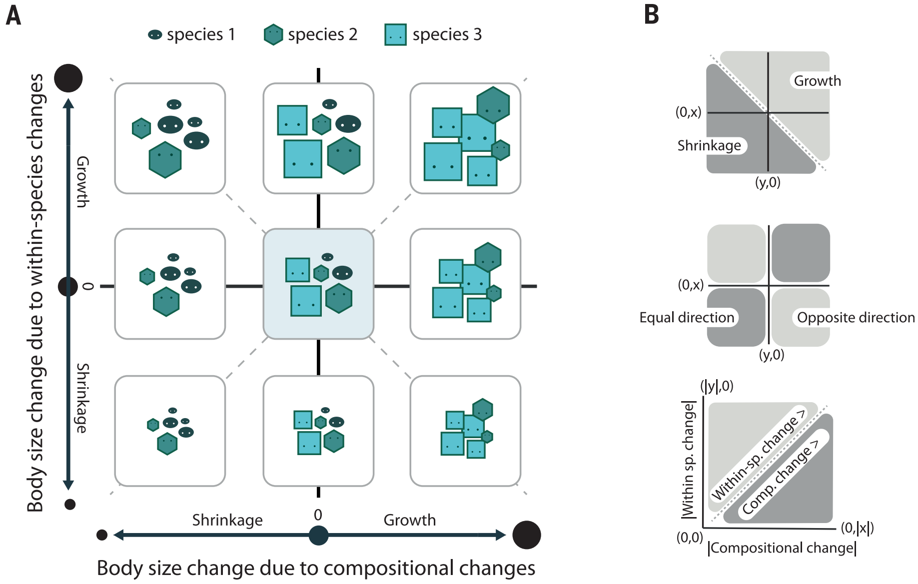


Fig. 1. Components of temporal changes in mean body size. (**A**) Shifts in mean assemblage-level body size can occur due to within-species changes (vertical axis), compositional changes (horizontal axis), or a combination of both components, displayed as change between two time points, from time 1 to time 2. Boxes represent assemblages made up of individual organisms (icons), with different colors and shapes representing different species. Icon size represents the body size of an individual within each species. The body size distribution at time 1 is shown in the middle (i.e., the example of no change from time 1 to time 2), with different change outcomes for time 2 shown in the other cartoons. Note that the vertical placement (axis) represents the within-species (population or intraspecific) changes through time in mean body size. This could be a mixture of increases due to smaller individuals growing larger or being replaced by larger individuals (shown in the right boxes) and decreases in the average size of individuals (shown in the left boxes). The horizontal placement (axis) indicates change in mean body size resulting from the gain or loss of species (compositional turnover) or a change in the relative abundance of the species present in an assemblage (even without local extinction or immigration of species). (**B**) The two components of body size change can reinforce or counteract each other. Assemblage body size change is most pronounced when both components operate in the same direction (toward either shrinking or increasing body size), such that the covariance between compositional and within-species changes is positive. When only one component is involved (i.e., change in one axis but not the other), body size change tends to be lower, and with negative covariance, it is possible to have change in one component cancel out change in the other (middle). If the components counteract each other, then the overall direction of change will depend on which component shows the higher absolute effect (contribution).

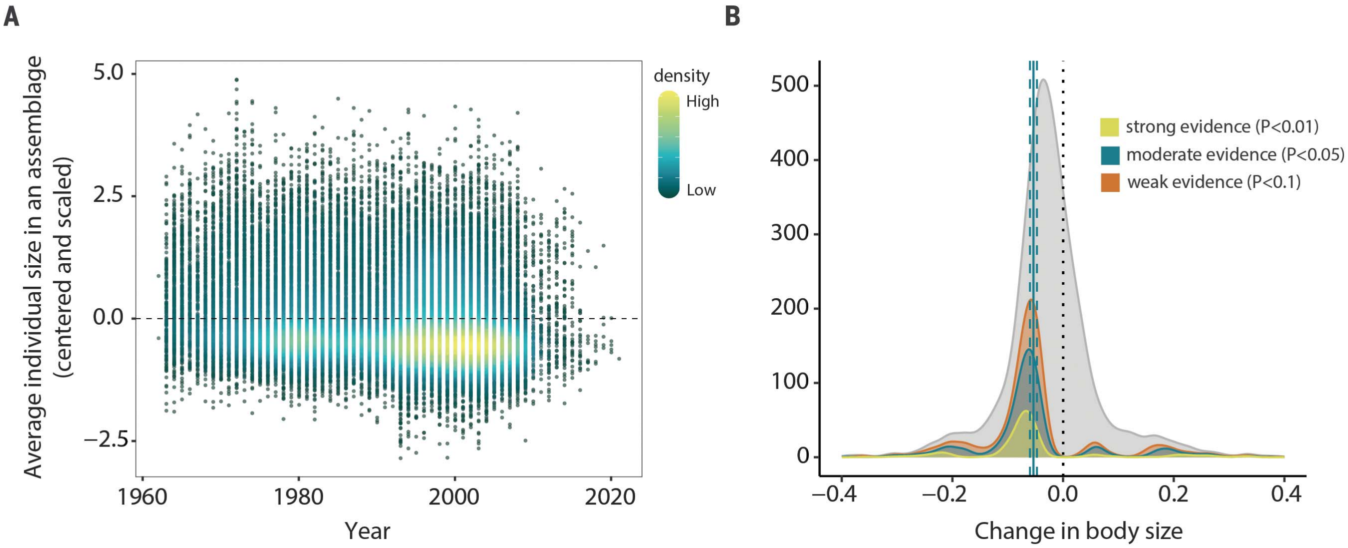


Fig. 2. Changes in mean body size across the 5025 assemblages. (**A**) Average individual size across the full set of assemblage time series. Each point is colored by density break, with colder colors indicating lower densities. (**B**) Density plots of the distribution of slopes of change in average individual body size. The full set of 5025 assemblage time series is shown in light gray. Yellow, blue, and orange represent, respectively, the subset of assemblages for which strong evidence ($P < 0.01$), moderate evidence ($P < 0.05$), and weak evidence ($P < 0.1$) of change was detected when testing slopes against 0. Dotted lines show slope of 0, and blue dashed lines show the mean slope across the blue data (traditional significance value) and the respective 90% credible interval.

in average body size and one-third increased. In fact, both components of body size change were present in the overwhelming majority of assemblages (96.4%), with the magnitude of compositional change being greater than that of within-species change in 72% of assemblages. Although compositional and within-species change often occurred in the same direction (58.8% of assemblages), we found counteracting effects in 41.2% of all assemblages. For example, of the 3415 assemblages showing within-species decreases in body size, ~35% had increases in body size associated with compositional change (within-species change <0 and compositional change >0; Fig. 3). This substantial variation in magnitude and direction of the two components of body size change and their interactions suggests that both components need to be considered when assessing change in body size. Although duration and time period of the time series vary, our results

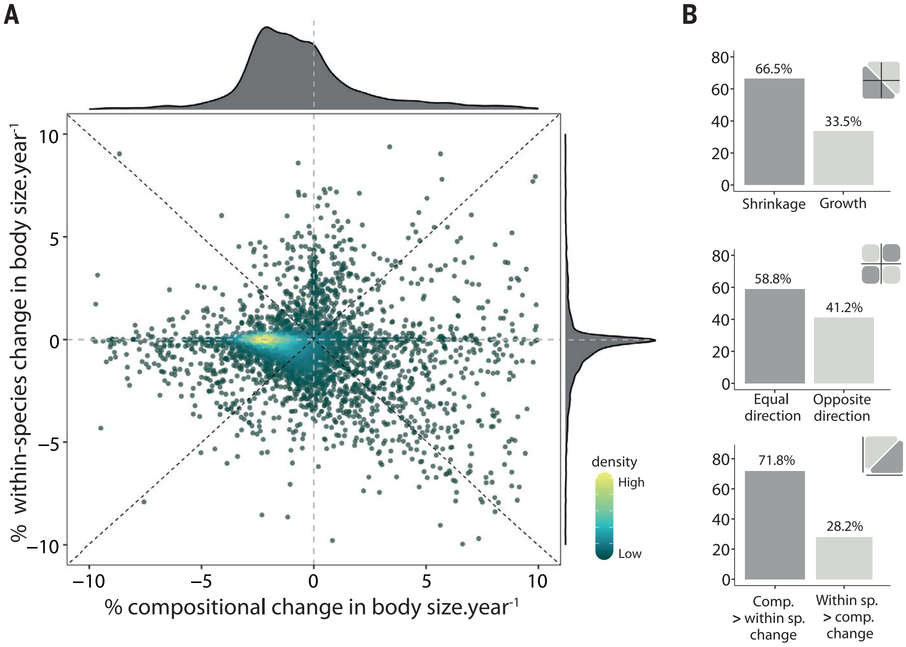


Fig. 3. Compositional versus within-species body size change through time in 5025 assemblages representing 4292 species of fish, plants, invertebrates, mammals, herpetofauna, and marine benthic organisms. (**A**) Relationship between population-level (i.e., within-species) changes and assemblage-level (i.e., compositional) changes. Both axes show percent changes standardized by the number of years between the first and last year sample of the assemblage (duration); assemblages (points) are colored by density break (with colder colors indicating lower densities). Dashed lines show $x = 0$, $y = 0$, $x = y$, and $y = -x$. (**B**) Frequency distributions (in percentage) of the number of assemblages (n = 5025) in the different scenarios depicted in Fig. 1B. Assemblages with a percentage change per year higher than 10% (n = 517; see figs. S2 and S7) are not shown in (A) but are included in (B).

are robust to the length of the time series, the start and end dates, as well as intermediate states (pairwise comparisons among years; see sensitivity analysis in figs. S5 and S6).

We found that trends in body size change over time vary among taxa, realm, and latitude (Fig. 4). Our confidence in estimates of body size change is highest for the most well-represented taxon in our dataset, marine fish, which show a particularly evident decrease in body size (Fig. 4A). Among other taxa, the number of available time series is lower and body size change trends are more variable. In fact, when fish are removed from our analysis, neither increases nor decreases dominate the overall body size change trends across the remaining dataset (fig. S3C). Among nonfish assemblages (e.g., benthos, mainly marine invertebrates and plants) the role of within-species and compositional changes is also more variable, but where the patterns of within-species are stronger (492 of 1116 nonfish time series), there is a tendency toward increasing body size (57% of assemblages), counteracting the tendency observed in fish assemblages (Fig. 4A). When we compared overlapping data (assemblages and species) with an extended dataset that uses species' average body size estimates taken from trait databases (fig. S2), we found marked consistency for both fish and nonfish assemblages (fig. S3, E to H). Nevertheless, we maintain that considering both axes of variation (compositional and within-species) is crucial to avoiding potentially misleading conclusions that arise when the two components change in opposite directions. For instance, when we estimated global trends in body size change across all available datasets by fitting a Bayesian mixed-effects model, we did not detect any clear pattern of change (with or without fish; figs. S9 to S11) (*19*). This was true regardless of whether we used the same data as in our main analysis that directly measured body size trends or the full extended dataset, which includes 20,173 assemblage time series with body size inferred from trait databases [figs. S2 and S3; see also (*14*)]. This extended dataset also highlights that neither increases nor decreases dominate mean body size change except in marine fish, even when there are few more (substantially more for birds) available time series for other taxa (fig. S12). This result emphasizes that we should take caution against extrapolating or overinterpreting trait changes across taxa, particularly when data on within-species change are not available. More data are needed to determine the prevalence of body size change through time in nonfish taxa.

Despite caveats, we consistently detected the signal of shrinking body size in marine fish for both types of body size data regardless of whether we used ordinary least-squares slope estimates or the Price equation approach. Our observation of shrinking among marine fish assemblages aligns with previous evidence (*25*–*27*). For marine fishes, these changes are often linked to the selective exploitation of large-bodied individuals by humans (*25*), to warming (*26*), and/or to decreased resource availability (*27*). Furthermore, disturbances and selective removal of larger individuals affect the age and size structure of populations, fish or otherwise, as well as the genetic structure within populations [e.g., plants (*28*)]. Such responses may be particularly prevalent among fish because of the widespread effects of overexploitation. Across assemblages, it is likely that combinations of these drivers result in the high variability in trends and prevalence of counteracting effects that we observed in these time series.

Collectively, our analyses reveal that both within-species and compositional changes combine to create high variability in the observed outcomes of assemblage-level body size changes through time. These findings highlight the importance of considering the separate and interactive effects of compositional and within-species body size change. Specifically, the community context is necessary to understand within-species change. For example, removing top predators (often the larger–body size individuals in an assemblage) can trigger mesopredator release, which alters assemblage size structure and composition (*29*). Although it is possible that some of these dynamics are missed if predators and prey belong to different sampled assemblages, the assemblage level is where regulation will play out, for example, in the context of ecological carrying capacity (*30*).

The selection forces acting on body size are varied and have heterogeneous distributions in space and time. By partitioning body size change into within-species and compositional change, as we have done here, we can begin to explain the wide variation in body size change patterns through time found in the literature. For example, global warming is simultaneously selecting for smaller body size (for metabolic reasons), affecting species' phenology and causing range shifts (*2*). Global warming and species phenology effects can best be seen in within-species changes, whereas range shifts induce compositional change. The net result of these processes will depend on the environmental context. For example, in the Arctic tundra, warming promotes larger shrubs (*6*) because species from warmer areas are expanding their ranges and there are longer growing seasons. By contrast, warming is associated with smaller fish in the North Sea (*9*), although selective harvesting and exploitation are likely also contributing

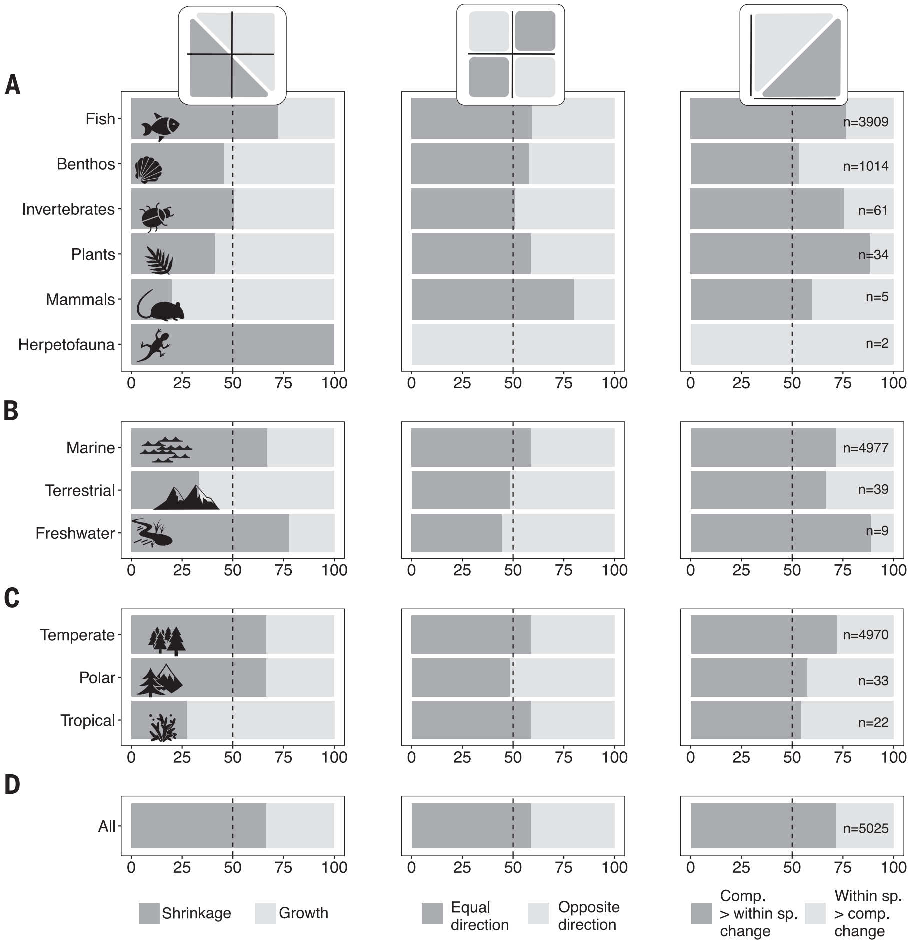


Fig. 4. Patterns of temporal body size change vary across taxa (A), realms (B), climates (C), and the globe (D). Plots show the frequency distributions (as percentages) of the number of assemblages across different groups for each scenario depicted in Fig. 1B. Dashed lines mark the 50% threshold.

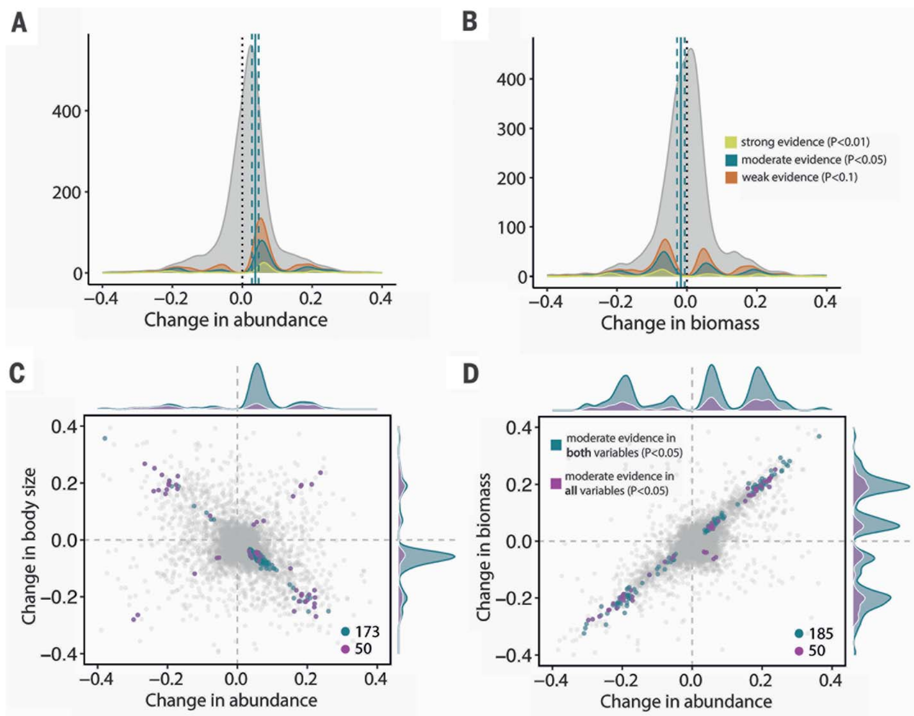


Fig. 5. Changes in assemblage abundance, biomass, and body size. (**A** and **B**) Density plots of the distribution of slopes of change in total abundance of individuals (of all species) (A) and change in total biomass in an assemblage as a function of time for the same assemblages as shown in Fig. 3 (B). The full set of 5025 assemblage time series is shown in light gray. Yellow, blue, and orange represent, respectively, the subset of assemblages for which strong evidence ($P < 0.01$), moderate evidence ($P < 0.05$), and weak evidence ($P < 0.1$) of change was detected when testing slopes against 0. Dotted lines show slope of 0, and blue dashed lines show the mean slope across the blue data (traditional significance value) and the respective 90% credible interval. (**C** and **D**) Bottom panels show the different relationships between variables. Only assemblages for which strong or moderate evidence ($P < 0.05$) were detected for both variables plotted are shown in blue, and purple highlights the assemblages for which significant changes through time were detected in all three variables ($n = 50$); all remaining assemblages are shown in light gray. (C) Change in average body size as a function of abundance changes (note that 79% of the blue dots are in the quadrant where abundance increases and body size decreases). (D) Change in biomass as a function of abundance changes.

to this change (*31*). By considering both within-species and compositional changes in individual-level body size alongside changes in relative abundance, future research should be able to better elucidate the mechanisms involved in how body size is changing through time.

Relationships among changes in body size, abundance, and biomass

Body size is usually tightly linked to abundance (*32*) through both metabolic (*33*) and trophic (*34*, *35*) processes. This relationship can have implications for assemblage biomass mediated by a trade-off between size and abundance (*36*). Thus, we further investigated whether changes in body size are associated with changes in assemblage abundance, biomass, or both (Fig. 5 and fig. S13). We found that abundance has, on average, slightly increased through time across assemblages, whereas the overall change in assemblage biomass is indistinguishable from zero (Fig. 5). Previously, no relationship (*14*) or complex relationships (*37*) have been found between body size and abundance changes, although for invertebrates, such a relationship is often negative (*38*). Although our results confirm that the relationship between abundance change and body size change is complex and variable, there are signs that the overall reduction in body size is being counteracted by increasing overall abundance (Fig. 5, A and C). Such trade-offs between abundance and body size are often expected (*32*) and affect ecosystem metabolic rate and function (*24*, *39*). We detected a strong positive covariance between change in biomass and abundance (Fig. 5D) but much weaker covariance between abundance and body size, with the strongest trends among these two variables tending to negative covariance (Fig. 5C). These patterns in covariance are robust to removing fish time series from the analysis despite the fact that overall change in body size is not detected in that case (fig. S14). In fact, 79% of the assemblages with detectable trends in both variables have abundance increases and body size decreases. These patterns suggest that assemblage body size, abundance, and biomass are linked, and that change in one has implications for change in the others. There is evidence of widespread regulation of assemblage-level variables (species richness and abundance) in which assemblages tend to return to previous levels after disturbances (*40*). The lack of a clear directional trend in biomass in our study suggests that it may be more tightly regulated than body size and abundance, which may cause trade-offs in change of the latter two variables.

Conclusions

We found evidence of widespread body size shrinking through time as a result of both population- and community-level changes despite substantial variation and overall stable assemblage-level biomass. Not all taxa contributed equally to the observed changes that we report. We found the most widespread declines among fish assemblages but a greater balance of increases and declines in other taxa. Body size is an easily measured, integrative, and important morphological trait that scales with many ecological characteristics of organisms and ecosystems, such as demographic rates, metabolism, and resource requirements (*41*, *42*). We reiterate pleas for more regular monitoring of body size (*43*), especially for taxa other than marine fish and ideally in conjunction with abundance estimates. Future research could focus on the implications of body size changes for ecosystem functions. For instance, cascading food web effects of shrinking body size could negatively affect human nutrition and associated economics [e.g., by affecting crop plants and protein sources such as fish (*44*)]. Moreover, shrinking body size through compositional change is likely to bring changes in other traits and therefore trigger additional impacts on ecosystem functioning (*8*). Our study suggests that the ubiquitous turnover in biodiversity composition currently unfolding (*11*, *12*) is a profound reshuffling of not only species but also key characteristics of living organisms.

REFERENCES AND NOTES

1. B. J. Enquist, A. J. Abraham, M. B. J. Harfoot, Y. Malhi, C. E. Doughty, *Nat. Commun.* **11**, 699 (2020).
2. J. L. Gardner, A. Peters, M. R. Kearney, L. Joseph, R. Heinsohn, *Trends Ecol. Evol.* **26**, 285–291 (2011).
3. I. A. Hatton, R. F. Heneghan, Y. M. Bar-On, E. D. Galbraith, *Sci. Adv.* **7**, eabh3732 (2021).
4. M. Heino, B. Díaz Pauli, U. Dieckmann, *Annu. Rev. Ecol. Evol. Syst.* **46**, 461–480 (2015).
5. F. A. Smith, R. E. Elliott Smith, S. K. Lyons, J. L. Payne, *Science* **360**, 310–313 (2018).
6. A. D. Bjorkman *et al*., *Nature* **562**, 57–62 (2018).
7. W. J. Ripple *et al*., *Conserv. Lett.* **12**, e12627 (2019).
8. I. Bartomeus *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4656–4660 (2013).
9. M. Daufresne, K. Lengfellner, U. Sommer, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 12788–12793 (2009).
10. F. He *et al*., *Glob. Chang. Biol.* **25**, 3883–3892 (2019).
11. S. A. Blowes *et al*., *Science* **366**, 339–345 (2019).
12. M. Dornelas *et al*., *Science* **344**, 296–299 (2014).
13. J. Dembitzer, R. Barkai, M. Ben-Dor, S. Meiri, *Quat. Sci. Rev.* **276**, 107316 (2022).
14. J. C. D. Terry, J. D. O'Sullivan, A. G. Rossberg, *Nat. Ecol. Evol.* **6**, 140–144 (2022).
15. H.-F. Wang, M. Xu, *Ecol. Evol.* **7**, 10815–10828 (2017).
16. T. Dayan, D. Simberloff, *Mammal Rev.* **28**, 99–124 (1998).
17. M. Dornelas *et al*., *Glob. Ecol. Biogeogr.* **27**, 760–786 (2018).
18. T. Carroll, F. Cardou, M. Dornelas, C. D. Thomas, M. Vellend, *Glob. Chang. Biol.* **29**, 3525–3538 (2023).
19. Materials and methods are available as supplementary materials.
20. S. A. Frank, *J. Evol. Biol.* **25**, 1002–1019 (2012).
21. A. Gardner, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190361 (2020).
22. G. R. Price, *Nature* **227**, 520–521 (1970).
23. M. Vellend, *Q. Rev. Biol.* **85**, 183–206 (2010).
24. P. Pillai, T. C. Gouhier, *Ecology* **100**, e02645 (2019).
25. N. E. Bosch *et al*., *Conserv. Biol.* **36**, e13807 (2022).
26. J. L. Gardner *et al*., *J. Avian Biol.* **45**, 529–535 (2014).
27. J. A. Sheridan, D. Bickford, *Nat. Clim. Chang.* **1**, 401–406 (2011).
28. D. B. Lindenmayer, W. F. Laurance, J. F. Franklin, *Science* **338**, 1305–1306 (2012).
29. T. M. Newsome *et al*., *Nat. Commun.* **8**, 15469 (2017).
30. J. H. Brown, S. K. M. Ernest, J. M. Parody, J. P. Haskell, *Oecologia* **126**, 321–332 (2001).
31. S. Jennings *et al*., *Mar. Biol.* **141**, 1085–1097 (2002).
32. E. P. White, S. K. M. Ernest, A. J. Kerkhoff, B. J. Enquist, *Trends Ecol. Evol.* **22**, 323–330 (2007).
33. J. Damuth, *Nature* **290**, 699–700 (1981).
34. J. E. Cohen, T. Jonsson, S. R. Carpenter, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1781–1786 (2003).
35. J. L. Blanchard *et al*., *J. Anim. Ecol.* **78**, 270–280 (2009).
36. E. P. White, S. K. M. Ernest, K. M. Thibault, *Am. Nat.* **164**, 670–676 (2004).
37. M. J. Genner *et al*., *Glob. Chang. Biol.* **16**, 517–527 (2010).
38. M. A. K. Gillespie, T. Birkemoe, A. Sverdrup-Thygeson, *Ecol. Evol.* **7**, 1091–1100 (2017).
39. J. H. Brown, J. F. Gillooly, A. P. Allen, V. M. Savage, G. B. West, *Ecology* **85**, 1771–1789 (2004).
40. N. J. Gotelli *et al*., *Sci. Adv.* **3**, e1700315 (2017).
41. M. Flannery, *Am. Biol. Teach.* **51**, 122–125 (1989).
42. M. Tseng *et al*., *J. Anim. Ecol.* **87**, 647–659 (2018).
43. G. Woodward *et al*., *Trends Ecol. Evol.* **20**, 402–409 (2005).
44. A. Audzijonyte, A. Kuparinen, R. Gorton, E. A. Fulton, *Biol. Lett.* **9**, 20121103 (2013).
45. I. S. Martins *et al*., Code and data for: Widespread shifts in body size within populations and assemblages, Zenodo (2023); https://doi.org/10.5281/zenodo.7969814.

ACKNOWLEDGMENTS

Funding: This work was supported by a Marie Sklodowska-Curie Actions Individual Fellowship (MSCA-IF); European Union Horizon 2020 (grant 894644 to I.S.M.); a German Research Foundation grant to the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig (DFG FZT-118 grant 202548816) sTeTra working group through sDiv, the Synthesis Centre of iDiv (F.S. and M.D.); a German Research Foundation grant to the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig (DFG FZT-118 grant 20254881 to S.A.B., J.M.C., N.E., R.v.K., and A.S.); a German Research Foundation grant to the Gottfried Wilhelm Leibniz Prize (grant Ei 862/29-1 to N.E.); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Coordination for the Improvement of Higher Education Personnel (CAPES process number 88881.129579/2016–01, finance code 001, to I.T.S.); the Leverhulme Trust (RPG-2019-402 to A.E.M. and M.D. and ECF-2021-512 to M.M.); Leverhulme Trust through the Leverhulme Centre for Anthropocene Biodiversity (RC-2018-021 to I.S.M. and M.D.); the European Union (ERC coralINT 101044975 to M.D.); a Schmidt Science Fellowship (G.N.D.); a Fisheries Society of the British Isles PhD Studentship (A.F-.E); the Knut och Alice Wallenberg Foundation (A.D.B.); the US Department of Agriculture (Hatch Grant MAFES 1011538 to B.M.); and the National Science Foundation (EPSCOR Track II Grant 2019470 to B.M.). **Author contributions:** Conceptualization: all authors; Data curation: I.S.M., V.B., C.F.Y.C., R.v.K., A.B., A.F.-E., and F.M.; Formal analysis: I.S.M.; Funding acquisition: I.S.M., F.S., and M.D.; Methodology: all authors; Project administration: I.S.M., M.D., F.S., and J.M.C.; Supervision: I.S.M., M.D., F.S., and J.M.C.; Visualization: I.S.M., M.D., F.S., C.F.Y.C., B.P., A.F-E., A.E.B., M.M., and S.A.B.; Writing – original draft: I.S.M., M.D., F.S., R.F. and J.M.C.; Writing – review and editing: all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The published BioTIME data can be accessed on Zenodo (*17*) or through the BioTIME website (https://biotime.st-andrews.ac.uk/). Links to the individual datasets are also provided in table S1. The selected data used in this article and the R scripts used to generate the main results of the study are archived online on Zenodo (*45*).

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg6006
Materials and Methods
Figs. S1 to S14
Table S1
References (*46*–*261*)
MDAR Reproducibility Checklist

Submitted 3 February 2023; accepted 10 August 2023
10.1126/science.adg6006

ORGANOMETALLICS

Oxidative addition of an alkyl halide to form a stable Cu(III) product

Yongrui Luo[1]†, Yuli Li[1]†, Jian Wu[1], Xiao-Song Xue[1]*, John F. Hartwig[2]*, Qilong Shen[1]*

The step that cleaves the carbon-halogen bond in copper-catalyzed cross-coupling reactions remains ill defined because of the multiple redox manifolds available to copper and the instability of the high-valent copper product formed. We report the oxidative addition of α-haloacetonitrile to ionic and neutral copper(I) complexes to form previously elusive but here fully characterized copper(III) complexes. The stability of these complexes stems from the strong Cu–CF_3 bond and the high barrier for C(CF_3)–C(CH_2CN) bond-forming reductive elimination. The mechanistic studies we performed suggest that oxidative addition to ionic and neutral copper(I) complexes proceeds by means of two different pathways: an S_N2-type substitution to the ionic complex and a halogen-atom transfer to the neutral complex. We observed a pronounced ligand acceleration of the oxidative addition, which correlates with that observed in the copper-catalyzed couplings of azoles, amines, or alkynes with alkyl electrophiles.

Copper-mediated cross-coupling reactions have become some of the most powerful methods for the construction of carbon-carbon (C–C) and C–heteroatom bonds (*1*–*4*). Early efforts in this field focused mainly on the coupling of sp^2-hybridized carbon electrophiles, and in recent years, research has expanded to encompass the mild coupling of sp^3-hybridized carbon electrophiles (*5*, *6*). Much progress has been made recently on copper-mediated or copper-catalyzed alkynylation (*7*–*9*), alkylation (*10*–*12*), arylation (*13*–*15*), and amination (*16*, *17*) of alkyl halides, providing an alternative and practical method for the installation of functional groups on alkyl chains and rings. Despite this expansion of the scope of the cross-coupling reactions of alkyl electrophiles that use copper, the mechanism of these reactions is poorly understood.

Previously, two distinct cycles were proposed for copper-catalyzed cross-coupling reactions of alkyl electrophiles. One cycle comprises a two-electron Cu(I)/Cu(III) manifold, and one comprises a stepwise Cu(I)/Cu(II) manifold involving initial single-electron transfer (SET) between the Cu(I) center and the alkyl electrophile to generate a Cu(II) intermediate and an alkyl radical, followed by transfer of the functional group from the resulting Cu(II) species to the alkyl radical (*18*–*22*) (Fig. 1A). Differentiation of these two pathways is challenging because the higher-valent copper intermediates in the reactions, particularly the putative Cu(III) intermediates, are highly reactive and typically elude detection (*23*–*26*).

[1]Key Laboratory of Organofluorine Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200032, PR China. [2]Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720, USA.
*Corresponding author. Email: xuexs@sioc.ac.cn (X-S.X.); jhartwig@berkeley.edu (J.F.H.); shenql@sioc.ac.cn (Q.S.)
†These authors contributed equally to this work.

Consequently, isolable, well-defined alkyl Cu(III) complexes are rare. In the early 2000s, seminal work from the groups of Bertz, Ogle (*27*–*29*), and Gschwind (*30*) demonstrated that reactions of alkyl iodides with ionic Gilman reagents generated Cu(III) species, which were characterized at a temperature below −93°C with rapid-injection nuclear magnetic resonance (RI-NMR) spectroscopy techniques (Fig. 1C). Yet, formation of the Cu(III) intermediates, even at this temperature, is fast, rendering studies on the mechanism of the reaction of alkyl halides with the Cu(I) species challenging.

To probe whether a Cu(III) intermediate could be generated from the reaction of an alkyl halide with a Cu(I) species and to determine how such an intermediate would form from an alkyl electrophile in a copper-mediated or copper-catalyzed cross-coupling, we reasoned that two requirements must be met: (i) An alkyl electrophile with a high reduction potential must be used so that formation of the Cu(III) species from the starting Cu(I) species is thermodynamically favorable, and (ii) the barrier for reductive elimination from the Cu(III) intermediate must be higher than that for the oxidative addition that forms the Cu(III) species (Fig. 1B).

The electron-withdrawing, strong-field trifluoromethyl ligand is known to stabilize both Cu(I) and high-valent Cu(III) metal centers as a result of π-back donation of electron density from the d orbitals on copper to the antibonding (σ*) orbitals of the C–F group or the contracted bonding (σ) orbitals on copper because of the high electronegativity of the fluorine (*31*). In the past three decades, a variety of well-defined trifluoromethyl Cu(III) complexes have been reported (*26*). Recent studies on the reactivities of these Cu(III) complexes showed that the barrier for reductive elimination to form a C(sp^3)–CF_3 bond from either the ionic Cu(III) complex $[Cu(CF_3)_3(alkyl)]^-$ or the five-coordinate neutral Cu(III) complex $[(bpy)Cu(CF_3)_2(Me)]$ is high (*32*–*34*). In addition, stable, well-characterized trifluoromethyl Cu(I) species—including $[CuCF_3]$ (*35*), which contains no additional ligands; $[(Phen)CuCF_3]$ (*36*), which contains a dinitrogen-donor ligand; $[(NHC)CuCF_3]$ (*37*), which contains an *N*-heterocyclic carbene (NHC) ligand; and $[Cu(CF_3)_2]^-$ (*38*, *39*), which is an ionic Cu(I) cuprate—were reported to react with aryl halides to give trifluoromethylarene products in good yields. In light of this prior work, we proposed that an appropriate trade-off between reactivity and stability of trifluoromethyl Cu(I) and Cu(III) complexes could meet the aforementioned requirements and provide an opportunity to investigate the mechanism of the reaction of alkyl halides with Cu(I) species to form stable Cu(III) products.

On the basis of the above rationale, we studied the reactions of alkyl halides with trifluoromethyl-Cu(I) complexes (Fig. 1). Stable $[Ph_4P]^+[Cu(CF_3)_2]^-$ and $[(bpy)Cu(CF_3)]$ (bpy, bipyridine) were chosen initially as the Cu(I) complexes that represent the ligandless "ate"-type Cu(I) complex and the neutral bipyridine-ligated Cu(I) complex, respectively. These complexes would allow us to probe the differences between the reactivity of two different types of Cu(I) complexes and the effect of the nitrogen ligand on the oxidative addition process. We chose α-haloacetonitrile XCH_2CN [in which X (halogen) is Cl, Br, or I] as the alkyl electrophile because XCH_2CN is more electrophilic than other alkyl halides, reducing the barrier for oxidative addition to Cu(I). In addition, because of the electron-withdrawing property of the cyano group, the barrier for C–C bond-forming reductive elimination from a $[(ligand)Cu^{III}(CF_3)_2(CH_2CN)]$ complex could be sufficiently high for the product to be observed directly and possibly isolated. We report the isolation of Cu(III) complexes from oxidative addition of haloacetonitrile to ionic and neutral trifluoromethyl Cu(I) complexes. Mechanistic investigation of these reactions, conducted with a combination of computational and experimental studies, shows that anionic and neutral complexes react with the same alkyl halide by means of distinct mechanisms.

Isolation and characterization of Cu(III) products

We initially studied the reactions of $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) or a neutral Cu(I) complex $[(bpy)Cu(CF_3)]$ (**1b**) with haloacetonitriles $ClCH_2CN$ (**2-Cl**), $BrCH_2CN$ (**2-Br**), and ICH_2CN (**2-I**) or alkyl tosylate $TsOCH_2CN$ (**2-OTs**) (Fig. 2A). The reaction of 2.0 equivalents of anion **1a** and bromide **2-Br** in dimethyl sulfoxide (DMSO) occurred smoothly at room temperature after 3.0 hours to give two previously unknown species in an approximate

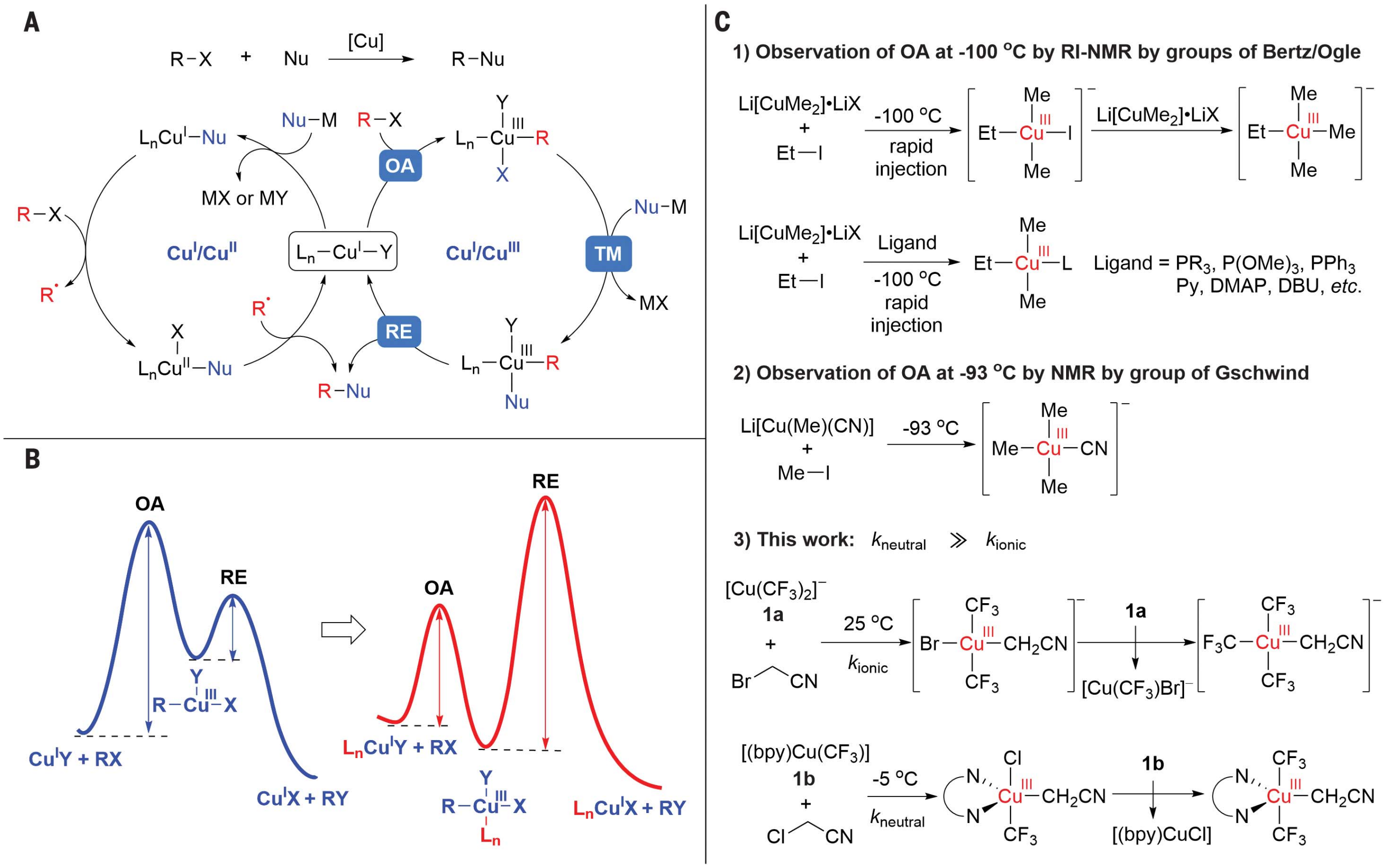


Fig. 1. Mechanism of Cu-mediated cross-coupling. (**A**) General mechanism for Cu-catalyzed cross-coupling reaction. (**B**) Strategy that inverted the barrier for oxidative addition (OA) and reductive elimination (RE) in the catalytic cycle for Cu-catalyzed cross-coupling. (**C**) State-of-the-art observations of oxidative addition of alkyl halide to Cu(I) species. M, metal or quasi-metal; X, halogen; Y, dummy ligand; TM, transmetalation.

1:1 ratio in quantitative yield, as determined by ^{19}F NMR spectroscopy. One species, corresponding to chemical shifts of −33.4 and −34.4 parts per million (ppm) in a 1:2 integral ratio in the ^{19}F NMR spectrum, was assigned as the tetracoordinate ionic Cu(III) complex $[Ph_4P]^+[Cu(CF_3)_3(CH_2CN)]^-$ (**3a**) (fig. S1). Cu(III) complex **3a** was stable enough to be isolated and characterized by ^1H, ^{19}F, and ^{31}P NMR spectroscopies and elemental analysis. The structure of complex **3a** was further confirmed by single-crystal x-ray diffraction, which revealed a typical square-planar geometry (Fig. 2C, top). The second complex, corresponding to a chemical shift at −27.0 ppm in the ^{19}F NMR spectrum, was assigned as a cuprate(I) species, $[Ph_4P]^+[Cu(CF_3)(Br)]^-$ (**1c**). We presume that complexes **3a** and **1c** were generated from transmetalation of the oxidative addition product $[Ph_4P]^+[Cu(CF_3)_2(CH_2CN)(Br)]^-$ with Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**). It was found that complex **1c** was much less reactive than complex **1a**. It reacted with 1.0 equivalent of bromide **2-Br** in DMSO, resulting in 30% conversion and 9% yield of **3a** after 3 hours at room temperature. Thus, the formation of bromocuprate **1c** does not affect the oxidative addition of **2-Br** to $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**). As did the reaction of bromide **2-Br**, the reaction of iodide **2-I** with cuprate **1a** occurred smoothly to give Cu(III) complex **3a** in more than 90% yield and $[Ph_4P]^+[Cu(CF_3)(I)]^-$ (**1d**) in 62% yield, respectively. By contrast, the reaction of chloride **2-Cl** with cuprate **1a** was much slower, and the starting materials remained intact even after 5 hours at room temperature (Fig. 2A). We also studied the reaction of tosylate **2-OTs** with **1a** and found that complex **1a** was fully converted within 3 hours at 25°C, but the yield for the formation of Cu(III) complex **3a** was much lower (15%) than those from the reactions of **2-Br** or **2-I** (Fig. 2A).

The reaction of neutral Cu(I) complex $[(bpy)Cu(CF_3)]$ (**1b**) with $BrCH_2CN$ (**2-Br**) occurred much faster than that of ionic Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**). The reaction of **1b** with bromide **2-Br** in *N*,*N*-dimethylformamide (DMF) generated *trans*-$[(bpy)Cu(CF_3)_2(CH_2CN)]$ (***trans*-4**) and *cis*-$[(bpy)Cu(CF_3)_2(CH_2CN)]$ (***cis*-4**) in 56 and 4% yields, as well as [(bpy)CuBr], respectively, after just 1 min at room temperature (Fig. 2B). Complex ***trans*-4** was fully characterized by ^1H and ^{19}F NMR spectroscopies, as well as elemental analysis.

X-ray diffraction of the single crystal of ***trans*-4** shows that it adopts a distorted square-pyramidal geometry. The H_2C–Cu–N angle to the apical nitrogen is 121.9°, and that to the equatorial nitrogen is 160.2°; the length of the N–Cu bond at the apical position (2.236 Å) is considerably longer than that of the N–Cu bond in the equatorial position (1.987 Å), indicating that the geometry is closer to a square-based pyramid than a trigonal bipyramid (Fig. 2C, bottom).

Reaction of complex **1b** with iodide **2-I** also occurred quickly to give ***trans*-4** and ***cis*-4** in 67 and 11% yields, respectively, under similar conditions (Fig. 2B). In contrast to cuprate **1a**, the neutral complex **1b** reacted with chloride **2-Cl** to give ***trans*-4** in 49% yield at room temperature after just 1 hour (Fig. 2B). These rapid rates indicate that the neutral Cu(I) complex **1b** is much more reactive toward the alkyl halide than is ionic Cu(I) complex **1a**, reflecting a sizable ligand acceleration (Fig. 2D). Such a phenomenon has previously been observed in various copper-mediated or -catalyzed cross-coupling reactions (*40*, *41*).

The large difference in reactivity of ionic and neutral Cu(I) complexes led us to conduct

reactions that would reveal the mechanisms by which the two complexes react with alkyl halides. The reaction of bromide **2-Br** with ionic Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) was sensitive to the polarity of the solvent. The reactions conducted in polar solvents, such as DMSO, *N,N*-dimethylacetamide (DMAc), and DMF, proceeded to 96, 87, and 76% conversion at room temperature over 5 hours and afforded Cu(III) complex **3a** in 93, 70, and 60% yields, respectively (table S1), whereas the same reactions in less-polar solvents, such as tetrahydrofuran (THF) or CH_2Cl_2, occurred more slowly (24 and 12% conversions after 5 hours at room temperature) to give complex **3a** in just 13 and 11% yields, respectively (Fig. 2E). By contrast, the reaction of chloride **2-Cl** or bromide **2-Br** with the neutral Cu(I) complex $[(bpy)Cu(CF_3)]$ (**1b**) was not sensitive to the polarity of the solvent. Reactions of complex **1b** with chloride **2-Cl** in the more-polar solvent DMF or the less-polar solvent THF occurred to full conversion after 1 hour at room temperature to afford Cu(III) complex ***trans*-4** in 59 and 82% yields, respectively (Fig. 2E). Quantitative assessment of the reaction of chloride **2-Cl** with complex **1b** in THF and DMF at −5°C showed that the rates of the two reactions are similar [$(3.02 \pm 0.10) \times 10^{-3}\ M^{-1}\cdot s^{-1}$ in THF and $(3.43 \pm 0.32) \times 10^{-3}\ M^{-1}\cdot s^{-1}$ in DMF,

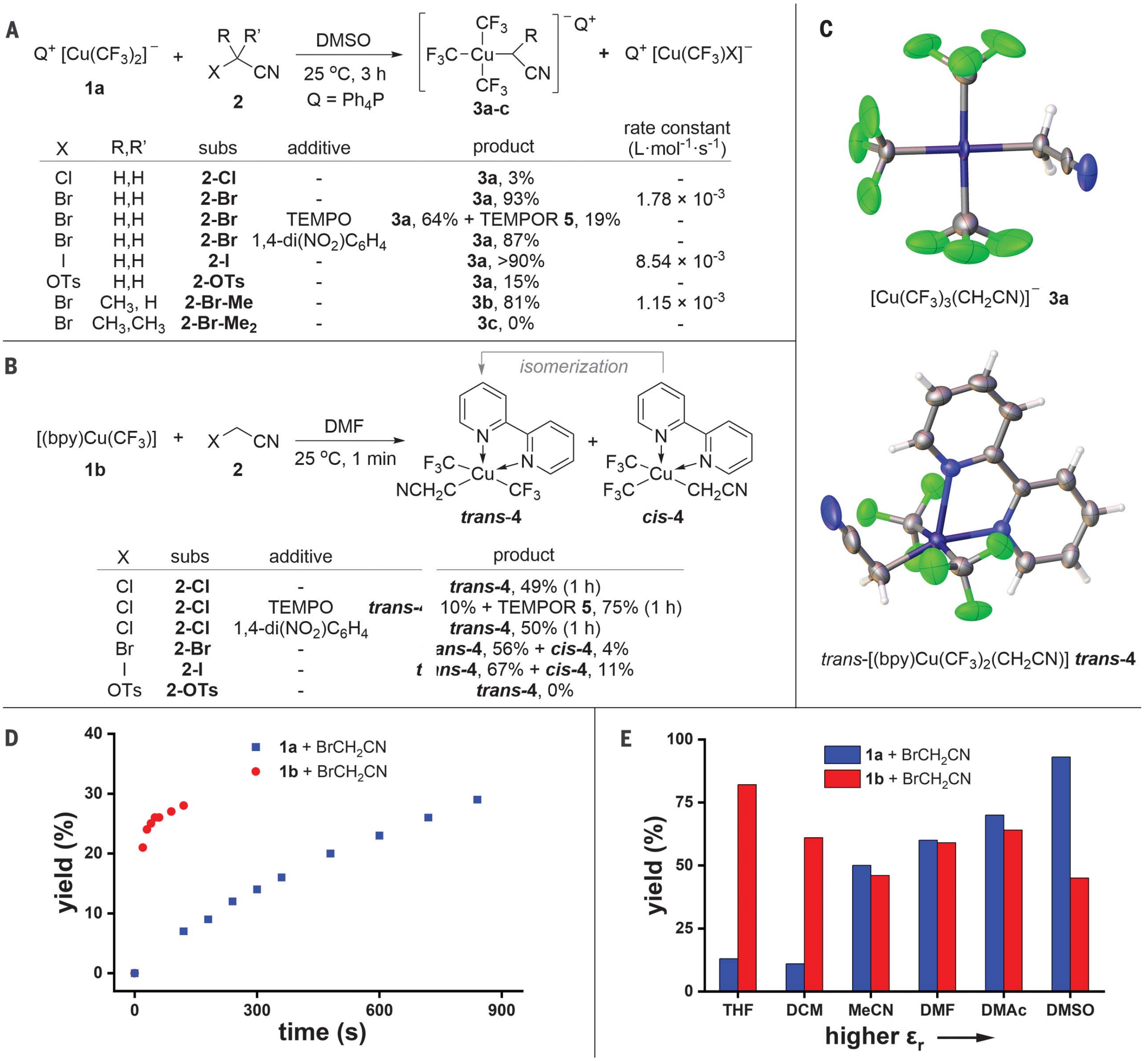


Fig. 2. Oxidative addition of alkyl halides to Cu(I) complexes. (**A**) Oxidative addition of XCH(R)CN with ionic Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**). (**B**) Oxidative addition of XCH(R)CN with $[(bpy)Cu(CF_3)]$ (**1b**). (**C**) ORTEP (Oak Ridge Thermal Ellipsoid Plot) diagrams of $[Ph_4P]^+[Cu(CF_3)_3(CH_2CN)]^-$ (**3a**) (countercation Ph_4P^+ is omitted for clarity) and *trans*-$[(bpy)Cu(CF_3)_2(CH_2CN)]$ (***trans*-4**). Ellipsoids are shown at the 50% level. (**D**) Reaction progress for ionic Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) with $BrCH_2CN$ (**2-Br**) at 298 K (blue) or neutral Cu(I) complex $[(bpy)Cu(CF_3)]$ (**1b**) with $BrCH_2CN$ (**2-Br**) at 298 K (red). (**E**) Solvent effect on the reactions of $BrCH_2CN$ with $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) (blue) or $[(bpy)Cu(CF_3)]$ (**1b**) (red). DCM, dichloromethane; MeCN, acetonitrile (methyl cyanide).

respectively (fig. S25)]. The different sensitivities of the ionic and neutral Cu(I) species to solvent polarity indicate that these Cu(I) complexes might react with the alkyl halide through different mechanisms.

Kinetic studies

To investigate the effect of the properties of the C–X bond of the haloacetonitriles on the reactions with cuprate $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) or neutral $[(bpy)Cu(CF_3)]$ (**1b**), we compared the rates of reactions of **1a** or **1b** with XCH_2CN quantitatively by monitoring the ^{19}F NMR signals corresponding to **1a** for reaction of complex **1a** and the signals corresponding to ***trans*-4** and ***cis*-4** for the reaction of complex **1b**. These studies showed that reactions of the ate complex are first order in complex **1a** and first order in bromide **2-Br** (Fig. 3A and figs. S3 to S6). The reaction of complex **1a** with iodide **2-I** [$(8.54 \pm 0.04) \times 10^{-3}\ M^{-1}\cdot s^{-1}$] was about five times as fast as that with bromide **2-Br** [$(1.78 \pm 0.03) \times 10^{-3}\ M^{-1}\cdot s^{-1}$ at 25°C]. Because complex **1a** reacted with chloride **2-Cl** slowly at room temperature (Figs. 2A and 3A), we studied the reaction at elevated temperature. At 60°C, the reaction occurred with a rate constant of $(1.99 \pm 0.14) \times 10^{-4}\ M^{-1}\cdot s^{-1}$. The rate constant for the reaction of complex **1a** with bromide **2-Br** at 60°C was estimated to be $4.05 \times 10^{-2}\ M^{-1}\cdot s^{-1}$ on the basis of the Eyring analysis in Fig. 3C, which is roughly 200 times as fast as that with chloride **2-Cl**.

The rates of the reactions of the halides with neutral complex **1b** were measured with ^{19}F NMR spectroscopy below room temperature because of their high rate. The reaction of bromide **2-Br** with $[(bpy)Cu(CF_3)]$ (**1b**) at –30°C proceeded to full conversion after 20 min. Kinetic studies showed that this reaction is first order in both reactants and that the rate constant at –30°C is $(2.63 \pm 0.05) \times 10^{-2}\ M^{-1}\cdot s^{-1}$. At this temperature after 5 min, the reaction of chloride **2-Cl** with $[(bpy)Cu(CF_3)]$ (**1b**) proceeded to <1% conversion (Fig. 3B). However, the reaction of chloride **2-Cl** with complex **1b** at –5°C occurred with a rate constant of $(3.43 \pm 0.32) \times 10^{-3}\ M^{-1}\cdot s^{-1}$. According to the Eyring analysis shown in Fig. 3C, the rate constant for reaction of complex **1b** with bromide **2-Br** at –5°C was estimated to be $0.37\ M^{-1}\cdot s^{-1}$, which is roughly 100 times greater than that of the reaction of complex **1b** with chloride **2-Cl.** The rate dependence of these reactions of the ionic and neutral Cu(I) species on the strength of the C–X bond and on the leaving group ability of the alkyl halides is consistent with typical Cu(I)-catalyzed cross-coupling reactions.

To evaluate the effect of the steric properties of the alkyl bromide on the reaction of the Cu(I) complex, we studied the reaction of Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) with the secondary alkyl halide, 2-bromopropionitrile **2-Br-Me**, and the tertiary alkyl halide, 2-methyl-2-bromopropionitrile **2-Br-Me$_2$** (Fig. 2A). Reaction of cuprate **1a** with 2-bromopropionitrile **2-Br-Me** occurred smoothly at room temperature over 5 hours to give Cu(III) complex **3b** in 81% yield. Yet, this reaction of the more sterically hindered 2-bromopropionitrile **2-Br-Me** was slower than that of the less-hindered bromoacetonitrile **2-Br** [$(1.15 \pm 0.01) \times 10^{-3}\ M^{-1}\cdot s^{-1}$ versus $(1.78 \pm 0.03) \times 10^{-3}\ M^{-1}\cdot s^{-1}$]. To determine whether the reactions of complex **1a** occurred with inversion of configuration, we conducted the reaction of **1a** with optically active (*R*)-2-bromopropionitrile (*R*)-**2-Br-Me**, but the enantiomers of the resulting product

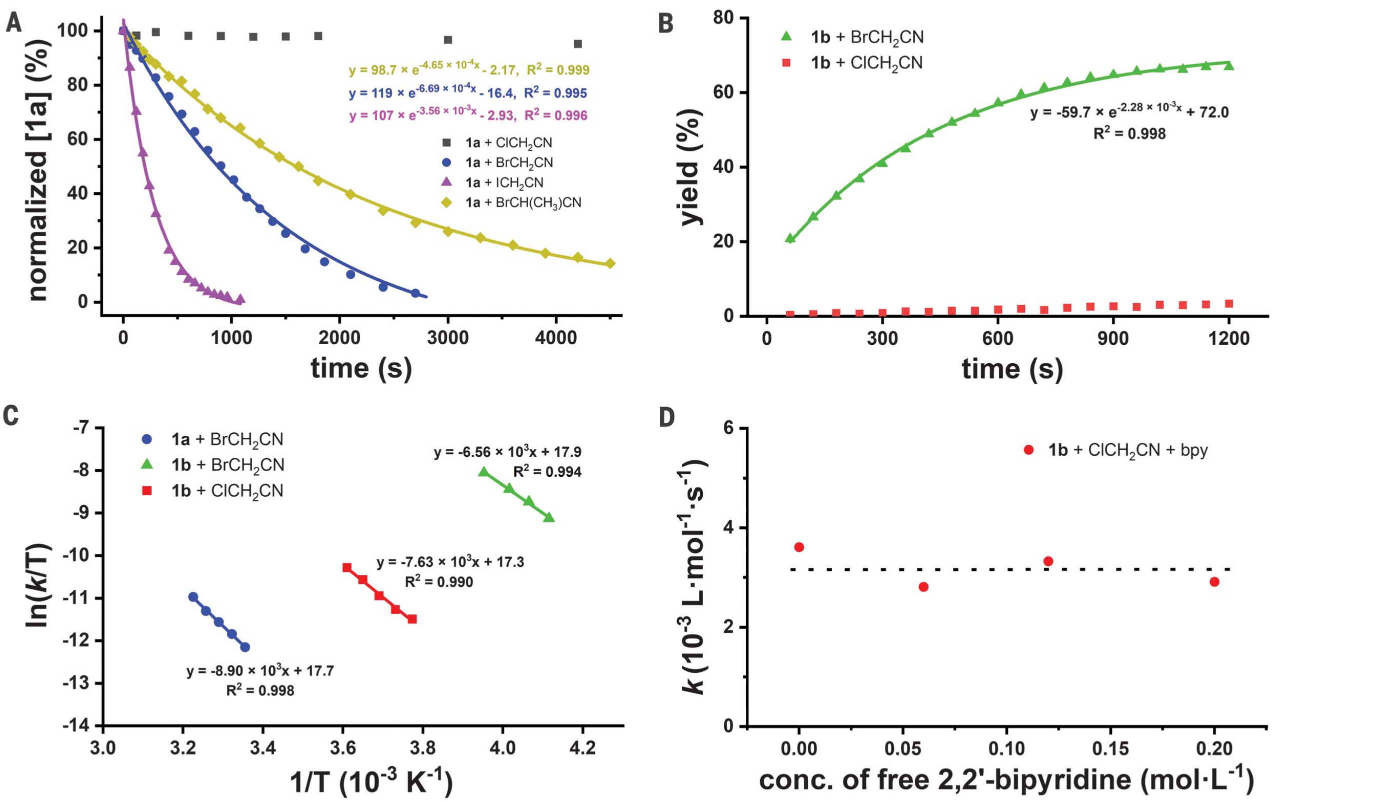


Fig. 3. Kinetic analysis. Kinetic data were fit to the expression of $[\mathbf{1a}]_t = [\mathbf{1a}]_0 e^{-k_{obs}} + c$ for (A) and $[\mathbf{4}]_t = A - Be^{-k_{obs}}$ for (B), in which t is time and k_{obs} are the apparent (observed) rate constants (pages S12 and S35 to S38 provide details about derivation of expression of corrected rate constant k_{corr} from k_{obs}). (**A**) Kinetic profiles of oxidative addition of XCH(R)CN (where X is Cl, Br, or I; and R is H or Me) with ionic Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) at 298 K. (**B**) Kinetic profiles of oxidative addition of XCH_2CN (where X is Cl or Br) with $[(bpy)Cu(CF_3)]$ (**1b**) at 243 K. (**C**) Eyring analysis of the temperature dependence of the rate constants of oxidative addition of $BrCH_2CN$ with $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) (blue), oxidative addition of $BrCH_2CN$ with $[(bpy)Cu(CF_3)]$ (**1b**) (green), and oxidative addition of $ClCH_2CN$ with $[(bpy)Cu(CF_3)]$ (**1b**) (red). (**D**) Effect of added free bipyridine on oxidative addition of $ClCH_2CN$ with $[(bpy)Cu(CF_3)]$ (**1b**) at 268 K.

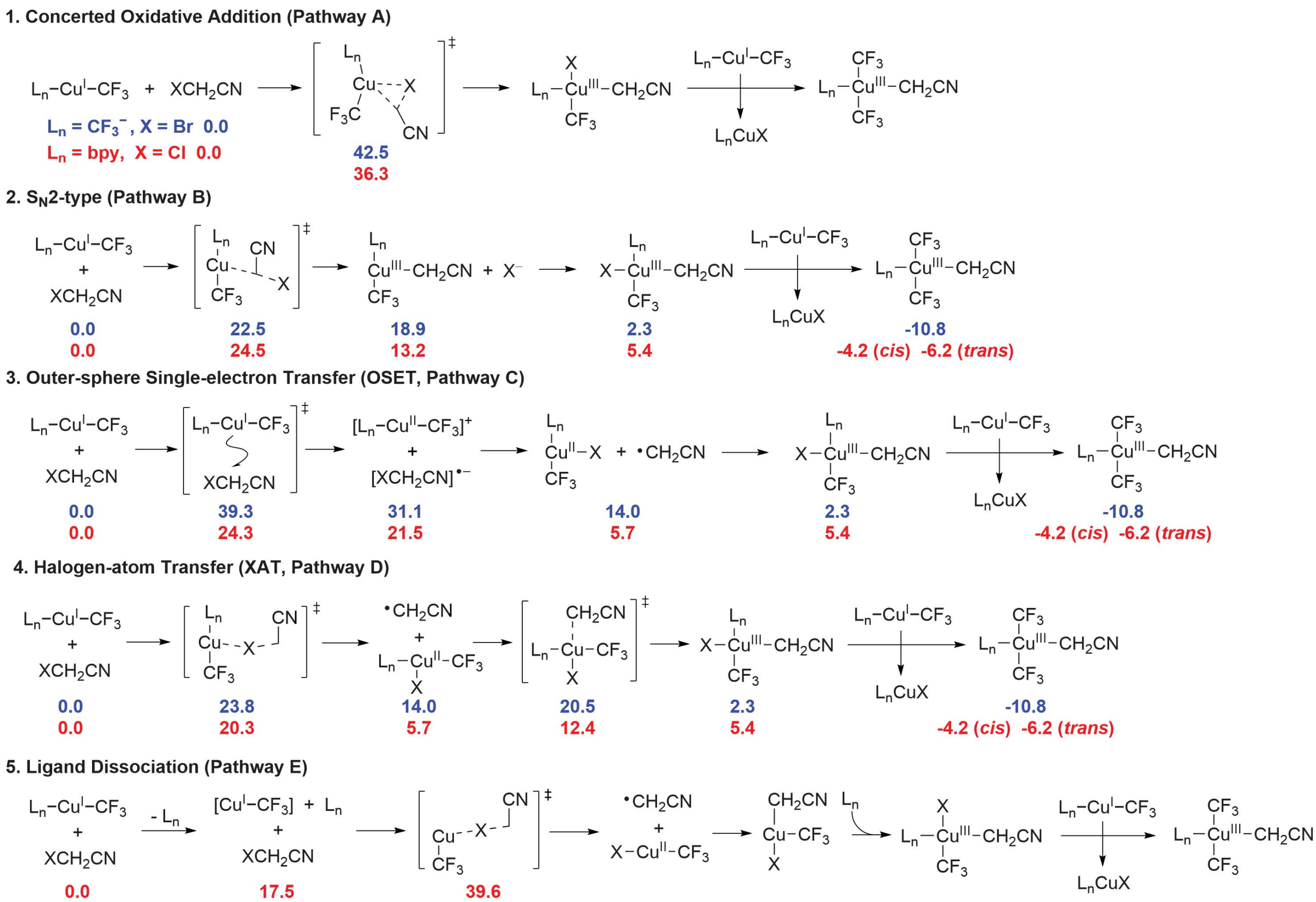


Fig. 4. Five proposed pathways for the oxidative addition of haloacetonitriles to ionic or neutral Cu(I) complexes and free energies of each species computed with DFT. The calculated activation free energies for the oxidative addition of $BrCH_2CN$(**2-Br**) to the ionic Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) in DMSO are given in blue, and those for oxidative addition of $ClCH_2CN$ (**2-Cl**) to the neutral $[(bpy)Cu(CF_3)]$ (**1b**) in DMF are given in red. The energies are in kilocalories per mole and indicate the relative free energies calculated at the PBE0-D3(BJ)/Def2-TZVP(SMD, solvent)//PBE0-D3(BJ)/Def2-SVP(SMD,solvent) level. SMD, solvation model density.

3b did not separate through chiral high-performance liquid chromatography (HPLC) column and only weakly absorbed in the ultraviolet, preventing assessment of the configuration of the product with chromatography or spectroscopy. The reaction of **1a** with the tertiary alkyl halide, 2-methyl-2-bromopropionitrile **2-Br-Me₂**, did not afford the corresponding Cu(III) complex from oxidative addition. Instead, Cu(III) complex $[Ph_4P]^+[Cu(CF_3)_4]^-$ and Cu(I) complex $[Ph_4P]^+[Cu(CF_3)(Br)]^-$ formed in 31 and 7% yields, respectively, as determined by ^{19}F NMR spectroscopy of the reaction mixture, as well as a few unidentified Cu(II) species, which were indicated by the color of the reaction mixture turning green. Analysis of the reaction mixture showed that isobutyronitrile **6** also formed through hydrodehalogenation. This observation suggests that the reaction forms the isobutyronitrile radical, which is too sterically hindered to combine with the trifluoromethyl Cu(II) species and, instead, abstracts a hydrogen atom from the solvent. The reaction of 2-chloropropionitrile **2-Cl-Me** with the neutral complex $[(bpy)Cu(CF_3)]$ (**1b**) occurred over 1 hour at 25°C to generate the corresponding Cu(III) complex ***trans*-4b** in 12% yield. This complex was characterized with ^{19}F NMR spectroscopy because it was too unstable to be isolated.

Inhibition studies

To probe whether the oxidative additions of XCH_2CN to Cu(I) complexes $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) and $[(bpy)Cu(CF_3)]$ (**1b**) occur by means of a radical intermediate and whether a potential radical would form through initial SET, we first studied the reaction in the presence of radical inhibitor 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) and in the presence of SET inhibitor 1,4-dinitrobenzene. The reaction of **1a** with $BrCH_2CN$ **2-Br** in the presence of 2.0 equivalents of TEMPO was complete after 3 hours at room temperature. The yield of this reaction (64%) was only 29% less than that in the absence of TEMPO, and only 19% of the radical adduct TEMPO–CH_2CN **5** was isolated (Fig. 2A). By contrast, TEMPO had a strong effect on the reaction of **1b** with $ClCH_2CN$ **2-Cl**, resulting in the formation of ***trans*-4** in 10% yield and TEMPO-CH_2CN **5** in 75% yield (Fig. 2B). SET inhibitor 1,4-dinitrobenzene did not noticeably affect the reaction of haloacetonitrile with either the ionic Cu(I) complex **1a** or the neutral Cu(I) complex **1b**. These results suggest that alkyl radicals are generated in the oxidative addition of **1a** or **1b** with bromoacetonitrile but that a SET process is unlikely to lead to the radical in either case.

Effect of ligand

The bipyridine ligand in complex **1b** could in principle dissociate from the metal center to create a less sterically hindered intermediate that reacts with the alkyl halide. If reaction of $ClCH_2CN$ with neutral Cu(I) complex $[(bpy)Cu(CF_3)]$ (**1b**) occurred through reversible dissociation

of the bipyridine, addition of free bipyridine to the reaction should substantially reduce the rate. The reactions in the presence of varying amounts of bipyridine (3.0, 6.0, and 10.0 equivalents versus **1b**) occurred with nearly equal rate constants [(2.81 ± 0.01) × 10^{-3} $M^{-1}{\cdot}s^{-1}$, (3.33 ± 0.21) × 10^{-3} $M^{-1}{\cdot}s^{-1}$, and (2.91 ± 0.12) × 10^{-3} $M^{-1}{\cdot}s^{-1}$, respectively] and were only slightly slower than the reaction in the absence of added 2,2′-bipyridine [(3.61 ± 0.08) × 10^{-3} $M^{-1}{\cdot}s^{-1}$] (Fig. 3D). These data imply that reversible dissociation of the ligand does not precede rate-limiting oxidative addition to **1b**.

Activation parameters

To determine the enthalpy and entropy of activation of both reactions, we studied the effect of the temperature on the rates. An Eyring plot of ln(k/T) versus $1/T$ (where k is the rate constant and T is temperature) for the reaction of ionic Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) with bromide **2-Br** between 25° and 37°C revealed an activation enthalpy, $\Delta H^\ddagger$, of 17.7 ± 0.4 kcal/mol and an activation entropy, $\Delta S^\ddagger$, of −12 ± 1 entropy units (e.u.) (Fig. 3C, blue). Likewise, an Eyring analysis of the reaction of neutral Cu(I) complex $[(bpy)Cu(CF_3)]$ (**1b**) with **2-Br** over a temperature range of −30° to −20°C revealed an activation enthalpy, $\Delta H^\ddagger$, of 13.0 ± 0.6 kcal/mol and a similar activation entropy, $\Delta S^\ddagger$, of −12 ± 2 e.u. (Fig. 3C, green). The activation entropies of the reactions of **1a** and of **1b** with **2-Br** were similar, but the activation enthalpy for the reaction of **2-Br** with **1b** was much less than that with **1a**. An Eyring analysis of the reaction of $[(bpy)Cu(CF_3)]$ (**1b**) with $ClCH_2CN$ (**2-Cl**) between −8° and 4°C revealed a $\Delta H^\ddagger$ of 15.2 ± 0.8 kcal/mol and a $\Delta S^\ddagger$ of −13 ± 3 e.u. (Fig. 3C, red), and this value for $\Delta H^\ddagger$ is similar to, or even slightly less than, that for oxidative addition of $BrCH_2CN$ (**2-Br**) to cuprate(I) complex $Ph_4P^+[Cu(CF_3)_2]^-$ (**1a**) (17.7 ± 0.4 kcal/mol), even though the C–Cl bond in $ClCH_2CN$ (**2-Cl**) (70.5 kcal/mol) is much stronger than the C–Br bond in $BrCH_2CN$ (**2-Br**) (56.8 kcal/mol) and the chloride is less polarizable and is a poorer leaving group than bromide (*42*).

To define the mechanism of the reaction of haloacetonitrile with the ionic or neutral Cu(I) species further, we performed density functional theory (DFT) calculations with the PBE0-D3(BJ) functional. On the basis of previous proposals for the mechanism of copper-mediated cross-coupling reactions (*18*–*22*) and the experimental results reported in this work, we proposed five different pathways that could account for the oxidative addition of haloacetonitrile to ionic and neutral Cu(I) complexes (Fig. 4). The first pathway (pathway A) involves an S_N2 step and is a common type of two-electron oxidative addition of alkyl halides to late transition metals, such as Pt (*43*) and Pd (*44*). In addition, such a pathway was proposed by Whitesides (*45*) and Pearson (*46*) for the reaction of lithium dialkylcuprates with alkyl halides (the Corey-Posner reaction) to generate a Cu(III) intermediate, which would then undergo fast reductive elimination to form the C–C bond in the product. In our case, reaction of **1a** through this mechanism would form Cu(III) species $[Cu^{III}(CF_3)_2(CH_2CN)]$ or $[Ph_4P]^+[Cu^{III}(CF_3)_2(X)(CH_2CN)]^-$, which would then undergo transmetalation with **1a** to generate $[Ph_4P]^+[Cu^{III}(CF_3)_3(CH_2CN)]^-$ (**3a**) and $[Cu(CF_3)X]^-$. Likewise, reaction with **1b** would give $[(bpy)Cu^{III}(CF_3)(CH_2CN)]^+$ or $[(bpy)Cu^{III}(CF_3)(X)(CH_2CN)]$, which would undergo transmetalation with **1b** to give $[(bpy)Cu^{III}(CF_3)_2(CH_2CN)]$ (***trans*-4** or ***cis*-4**) and [(bpy)CuX], respectively (details about transmetalation are provided in fig. S31). A second two-electron mechanism for oxidative addition would be a concerted addition of the C–X bond to the metal center (pathway B), and this step would be the reverse reaction of concerted reductive elimination from a high-valent Cu center. A third pathway for oxidative addition of alkyl halides to a Cu(I) center could occur through consecutive single-electron steps. One commonly proposed mechanism for oxidative addition through single-electron steps is outer-sphere SET (OSET; pathway C), which dominates the redox manifolds of first-row transition metals, including Fe (*47*) and Ni (*48*). A fourth pathway and an alternative mechanism for oxidative addition through single-electron steps is halogen-atom transfer (XAT; pathway D) (*49*, *50*). Ligated Cu(I) complexes are widely used in atom-transfer radical addition or polymerization (ATRA or ATRP) processes in which a XAT to Cu is involved (*51*). A fifth pathway, oxidative addition of the alkyl halide to Cu(I), could occur by initial ligand dissociation to generate a neutral, ligandless $[CuCF_3]$, which would undergo oxidative addition of the alkyl halide to generate a Cu(III) intermediate that would recoordinate the dative ligand and undergo transmetalation to give the final Cu(III) complex (pathway E) (*52*). The energy parameters for these pathways were computed and are shown in Fig. 4. In all five proposed mechanistic pathways, the formation of final product **3a** from the reaction of **1a** and the formation of ***trans*-4**/***cis*-4** from reaction of **1b** involve a transmetalation step after initial formation of a Cu(III) complex. The absence of observed intermediates before transmetalation and first-order rate behavior in the Cu(I) species (**1a** or **1b**) rule out both rate-limiting transmetalation for these reactions and the initial generation of the Cu(III) intermediate through reaction of two copper complexes.

Evidence for XAT and S_N2 pathways

On the basis of the experimental results and DFT calculations, we concluded that reaction of the ionic Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) with $BrCH_2CN$ (**2-Br**) likely proceeds by means of two different pathways, specifically the major fraction through an S_N2-type process (pathway A) and the minor fraction through a XAT process (pathway D). We also conclude that the reaction of the neutral Cu(I) complex $[(bpy)Cu(CF_3)]$ (**1b**) with $ClCH_2CN$ (**2-Cl**) likely proceeds exclusively through a XAT process (pathway D). The following analysis of our experimental and computational data lead to these conclusions.

The experimental observation that both reactions were partially or fully inhibited by the presence of the radical inhibitor TEMPO argues against a concerted pathway for oxidative addition of $BrCH_2CN$ (**2-Br**) to ionic Cu(I) complex $[Ph_4P]^+[Cu(CF_3)_2]^-$ (**1a**) or for oxidative addition of $ClCH_2CN$ (**2-Cl**) to neutral Cu(I) complex $[(bpy)Cu(CF_3)]$ (**1b**) (Fig. 4, pathway A). In addition, DFT calculations showed that the computed barrier (42.5 kcal/mol) for oxidative addition of $BrCH_2CN$ (**2-Br**) to complex **1a** through a concerted pathway (pathway A) is much greater than that of the S_N2-type pathway (pathway B; 22.5 kcal/mol) (Fig. 4) or XAT process (pathway D; 23.8 kcal/mol) (Fig. 4). Likewise, the computed barrier (36.3 kcal/mol) for oxidative addition of $ClCH_2CN$ (**2-Cl**) to complex **1b** by pathway A is about 16.0 kcal/mol greater than the lowest-energy alternative pathway, in this case the XAT process (pathway D; 20.3 kcal/mol).

The experimental observation that the addition of SET inhibitor 1,4-dinitrobenzene did not greatly affect either oxidative addition process argues against an OSET pathway (Fig. 4, pathway C). Moreover, the computed barriers for both reactions through an OSET pathway (39.3 kcal/mol and 24.3 kcal/mol, respectively) are 16.8 kcal/mol and 4.0 kcal/mol greater than the S_N2-type pathway A for reaction with complex **1a** or XAT pathway D for reaction with complex **1b**. Thus, both the experimental and computational data are inconsistent with reaction through an OSET pathway (Fig. 4, pathway C).

By contrast, the relative reactivity of the alkyl electrophiles toward the ionic Cu(I) complex **1a** of ICH_2CN > $BrCH_2CN$ >> $TsOCH_2CN$ >> $ClCH_2CN$ is consistent with an S_N2-type pathway B. In addition, the reaction of the secondary alkyl bromide 2-bromopropionitrile **2-Br-Me** with ionic Cu(I) complex **1a**, which is 1.5 times as slow as that of the less-hindered bromoacetonitrile **2-Br** [(1.15 ± 0.01) × 10^{-3} $M^{-1}{\cdot}s^{-1}$ versus (1.78 ± 0.03) × 10^{-3} $M^{-1}{\cdot}$ s^{-1}], is inconsistent with OSET or XAT pathways (Fig. 4, pathways C and D). The reaction of a secondary alkyl halide through an OSET or XAT pathway would be faster than that of the primary alkyl halide because of the greater stability of the secondary alkyl radical.

Also consistent with reaction through an S_N2 pathway B, the reactions of complex **1a**

with $BrCH_2CN$ (**2-Br**) in polar solvents were faster than those in less-polar solvents, and the more sterically hindered 2-bromopropionitrile **2-Br-Me** reacted more slowly than did **2-Br**. The generation of TEMPO–CH_2CN as a side product from the reaction of complex **1a** with $BrCH_2CN$ (**2-Br**) in the presence of TEMPO suggests that an alternative but minor pathway also occurs during the reaction of the cuprate **1a**. Consistent with this experimental observation, the barrier for a reaction of **1a** with $BrCH_2CN$ (**2-Br**) through a XAT pathway D (23.8 kcal/mol) computed with DFT was only 1.3 kcal/mol greater than that of the S_N2-type pathway B. Furthermore, the calculated barrier (22.5 kcal/mol for S_N2-type pathway B and 23.8 kcal/mol for XAT pathway D) is close to the corresponding experimentally observed activation energy (21.3 kcal/mol), which provides additional evidence to support the proposed mechanism for reaction of **1a** with haloacetonitrile.

Our data for reaction of the neutral copper complex **1b** are more consistent with XAT pathway D as the dominant mechanism. The considerable inhibiting effect of the addition of TEMPO on the reaction of chloroacetonitrile **2-Cl** with [(bpy)Cu(CF_3)] (**1b**) suggests that a free radical •CH_2CN is generated. The barrier for an OSET pathway C computed with DFT (24.3 kcal/mol) is much higher than that of the XAT process (pathway D, 20.3 kcal/mol). In addition, the calculated Gibbs free energy (ΔG; 20.3 kcal/mol) is close to the barrier (18.6 kcal/mol) determined experimentally, implying that a XAT process (pathway D) is the most likely pathway involving an alkyl radical for oxidative addition of alkyl halides to the neutral Cu(I) complex **1b**.

Our experimental and computational data are most consistent with reaction of the alkyl halide directly with [(bpy)Cu(CF_3)] (**1b**). The zero-order dependence on ligand is inconsistent with reversible ligand dissociation followed by oxidative addition (Fig. 4, pathway E). Furthermore, the computed ΔG for dissociation of bipyridine from complex **1b** (17.5 kcal/mol) is accessible, but the computed barrier for a XAT process (pathway D) from $ClCH_2CN$ to the resulting ligandless [$CuCF_3$] is an additional 22.1 kcal/mol. The combination of these energies is much greater than the computed $\Delta G^{\ddagger}$ (20.3 kcal/mol) for direct XAT to **1b** (pathway D).

To gain more insight into the oxidative addition of the haloacetonitrile to Cu(I) complex **1a** and **1b**, we conducted natural population analysis (NPA) on reactants **1a**, **1b**, **2-Br**, and **2-Cl**; transition states **TS-a-S$_N$2**, **TS-a-XAT** and **TS-b-XAT**; and products [$L_nCu^{III}(CF_3)(X)(CH_2CN)$] (where L_n is CF_3^- or bpy) (figs. S33 and S35). These studies showed that a small amount of positive charge accumulates on the copper in the transition states (+0.26 for **1a** versus +0.35 for [$Cu^{III}(CF_3)_2(Br)(CH_2CN)]^-$) during the reaction of **1a** with **2-Br** (*53*, *54*) and that the negative charge of the CF_3 moiety decreases considerably (−1.26 for **1a** versus −0.59 for [$Cu^{III}(CF_3)_2(Br)(CH_2CN)]^-$). Even though the change in charge to copper from starting complex to the intermediate after oxidative addition is small, the sum of the negative charge of the bromide and cyanomethyl ligands (−CH_2CN) is large (−0.76 e^-). The same trend was also observed for the reaction of **2-Cl** with **1b**. These changes in charge show that electron density flows from the complexes **1a** or **1b** overall to the haloacetonitriles during the reaction, thus demonstrating that reaction of haloacetonitrile to **1a** or **1b** can be considered an oxidative addition process.

Previous proposals for the mechanism of copper-catalyzed cross-coupling often involve a Cu(I)/Cu(II) redox cycle, on the basis of the experimental evidence for the presence of free radicals, which was typically deduced from quenching experiments with radical scavengers or from racemization or rearrangements of radical clock substrates. Our studies suggest that a free radical could be involved in the oxidative addition of an alkyl halide to Cu (I) to form a Cu(III) intermediate in these pathways through XAT. Oxidative addition of the C(sp^3)–X bond to a Cu(I) species is often considered to be the rate-limiting step of copper-catalyzed cross-coupling reactions of alkyl electrophiles, but studies of this elementary step alone are rare, largely because of the inherent instability of the copper intermediate. Thus, the example of oxidative addition of a C(sp^3)–X bond to Cu(I) in the current study may help develop more efficient copper-catalyzed cross-coupling reactions of alkyl electrophiles.

REFERENCES AND NOTES

1. I. Beletskaya, A. V. Cheprakov, *Coord. Chem. Rev.* **248**, 2337–2364 (2004).
2. G. Evano, N. Blanchard, M. Toumi, *Chem. Rev.* **108**, 3054–3131 (2008).
3. F. Monnier, M. Taillefer, *Angew. Chem. Int. Ed.* **48**, 6954–6971 (2009).
4. S. Bhunia, G. G. Pawar, S. V. Kumar, Y. Jiang, D. Ma, *Angew. Chem. Int. Ed.* **56**, 16136–16179 (2017).
5. L.-J. Cheng, N. P. Mankad, *Chem. Soc. Rev.* **49**, 8036–8064 (2020).
6. H. Zhou, Z.-L. Li, Q.-S. Gu, X.-Y. Liu, *ACS Catal.* **11**, 7978–7986 (2021).
7. X.-Y. Dong *et al.*, *Nat. Chem.* **11**, 1158–1166 (2019).
8. X. Mo, B. Chen, G. Zhang, *Angew. Chem. Int. Ed.* **59**, 13998–14002 (2020).
9. F. L. Wang *et al.*, *Nat. Chem.* **14**, 949–957 (2022).
10. E. J. Corey, G. H. Posner, *J. Am. Chem. Soc.* **89**, 3911–3912 (1967).
11. D. H. Burns, J. D. Miller, H.-K. Chan, M. O. Delaney, *J. Am. Chem. Soc.* **119**, 2125–2133 (1997).
12. G.-Z. Wang *et al.*, *Org. Lett.* **17**, 3682–3685 (2015).
13. C.-T. Yang, Z.-Q. Zhang, Y.-C. Liu, L. Liu, *Angew. Chem. Int. Ed.* **50**, 3904–3907 (2011).
14. X.-L. Su *et al.*, *Angew. Chem. Int. Ed.* **60**, 380–384 (2021).
15. C. Li, B. Chen, X. Ma, X. Mo, G. Zhang, *Angew. Chem. Int. Ed.* **60**, 2130–2134 (2021).
16. Q. M. Kainz *et al.*, *Science* **351**, 681–684 (2016).
17. C. Chen, J. C. Peters, G. C. Fu, *Nature* **596**, 250–256 (2021).
18. N. Yoshikai, E. Nakamura, *Chem. Rev.* **112**, 2339–2372 (2012).
19. C. Sambiagio, S. P. Marsden, A. J. Blacker, P. C. McGowan, *Chem. Soc. Rev.* **43**, 3525–3550 (2014).
20. S.-J. Li, Y. Lan, *Chem. Commun.* **56**, 6609–6619 (2020).
21. J. M. Ahn, T. S. Ratani, K. I. Hannoun, G. C. Fu, J. C. Peters, *J. Am. Chem. Soc.* **139**, 12716–12723 (2017).
22. H. Lee *et al.*, *J. Am. Chem. Soc.* **144**, 4114–4123 (2022).
23. A. J. Hickman, M. S. Sanford, *Nature* **484**, 177–185 (2012).
24. A. Casitas, X. Ribas, *Chem. Sci.* **4**, 2301–2318 (2013).
25. L. Liu, Z. Xi, *Chin. J. Chem.* **36**, 1213–1221 (2018).
26. H. Liu, Q. Shen, *Coord. Chem. Rev.* **442**, 213923 (2021).
27. S. H. Bertz, S. Cope, M. Murphy, C. A. Ogle, B. J. Taylor, *J. Am. Chem. Soc.* **129**, 7208–7209 (2007).
28. S. H. Bertz, S. Cope, D. Dorton, M. Murphy, C. A. Ogle, *Angew. Chem. Int. Ed.* **46**, 7082–7085 (2007).
29. E. R. Bartholomew *et al.*, *Chem. Commun.* (10): 1176–1177 (2008).
30. T. Gärtner, W. Henze, R. M. Gschwind, *J. Am. Chem. Soc.* **129**, 11362–11363 (2007).
31. P. J. Brothers, W. R. Roper, *Chem. Rev.* **88**, 1293–1326 (1988).
32. H. Shen *et al.*, *J. Am. Chem. Soc.* **139**, 9843–9846 (2017).
33. M. Paeth *et al.*, *J. Am. Chem. Soc.* **141**, 3153–3159 (2019).
34. S. Liu *et al.*, *J. Am. Chem. Soc.* **142**, 9785–9791 (2020).
35. A. Zanardi, M. A. Novikov, E. Martin, J. Benet-Buchholz, V. V. Grushin, *J. Am. Chem. Soc.* **133**, 20901–20913 (2011).
36. H. Morimoto, T. Tsubogo, N. D. Litvinas, J. F. Hartwig, *Angew. Chem. Int. Ed.* **50**, 3793–3798 (2011).
37. G. G. Dubinina, H. Furutachi, D. A. Vicic, *J. Am. Chem. Soc.* **130**, 8600–8601 (2008).
38. H. Liu, Q. Shen, *Org. Chem. Front.* **6**, 2324–2328 (2019).
39. S. Martínez de Salinas *et al.*, *Chem. Eur. J.* **25**, 9390–9394 (2019).
40. J. W. Tye, Z. Weng, A. M. Johns, C. D. Incarvito, J. F. Hartwig, *J. Am. Chem. Soc.* **130**, 9971–9983 (2008).
41. J. W. Tye, Z. Weng, R. Giri, J. F. Hartwig, *Angew. Chem. Int. Ed.* **49**, 2185–2189 (2010).
42. A. A. Isse, A. Gennaro, *J. Phys. Chem. A* **108**, 4180–4186 (2004).
43. L. M. Rendina, R. J. Puddephatt, *Chem. Rev.* **97**, 1735–1754 (1997).
44. J. K. Stille, K. S. Y. Lau, *J. Am. Chem. Soc.* **98**, 5841–5849 (1976).
45. G. M. Whitesides, W. F. Fischer Jr., J. San Filippo Jr., R. W. Bashe, H. O. House, *J. Am. Chem. Soc.* **91**, 4871–4882 (1969).
46. R. G. Pearson, C. D. Gregory, *J. Am. Chem. Soc.* **98**, 4098–4104 (1976).
47. J. D. Sears, P. G. N. Neate, M. L. Neidig, *J. Am. Chem. Soc.* **140**, 11872–11883 (2018).
48. G. D. Jones *et al.*, *J. Am. Chem. Soc.* **128**, 13175–13183 (2006).
49. J. B. Diccianni, J. Katigbak, C. Hu, T. Diao, *J. Am. Chem. Soc.* **141**, 1788–1796 (2019).
50. C. Fang *et al.*, *J. Am. Chem. Soc.* **141**, 7486–7497 (2019).
51. K. Matyjaszewski, J. Xia, *Chem. Rev.* **101**, 2921–2990 (2001).
52. P. Novák, A. Lishchynskyi, V. V. Grushin, *J. Am. Chem. Soc.* **134**, 16167–16170 (2012).
53. I. M. DiMucci *et al.*, *J. Am. Chem. Soc.* **141**, 18508–18520 (2019).
54. J. Shearer, D. Vasiliauskas, K. M. Lancaster, *Chem. Commun.* **59**, 98–101 (2022).

ACKNOWLEDGMENTS

Funding: Q.S. gratefully acknowledges financial support from the National Key Research and Development Program of China (2021YFF0701700) and the National Natural Science Foundation of China (22061160465). X-S.X. acknowledges financial support from the National Natural Science Foundation of China (22122104). J.F.H. acknowledges funding from the NIH (R35GM130387). Y. Luo thanks Syngenta for a PhD scholarship. **Author contributions:** Y. Luo performed the experiments and analyzed experimental data. Y. Li performed the DFT calculations. J.W. assisted in the kinetic studies. X.-S.X., J.F.H., and Q.S. directed the research. Y. Luo, J.F.H., and Q.S. wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data materials and availability:** Crystallographic data are available free

of charge from the Cambridge Crystallographic Data Centre (CCDC) under CCDC 2236874 (**3a**), 2236875 (**3b**), and 2236887 (***trans*-4**). All other data reported in this paper are available in the manuscript or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg9232
Materials and Methods
Figs. S1 to S36
Tables S1 to S15
References (*55*–*76*)

Submitted 26 May 2023; accepted 10 August 2023
10.1126/science.adg9232

ORGANOMETALLICS

Cross-coupling by a noncanonical mechanism involving the addition of aryl halide to Cu(II)

Connor P. Delaney[1], Eva Lin[1], Qinan Huang[1], Isaac F. Yu[1], Guodong Rao[2], Lizhi Tao[2], Ana Jed[1], Serena M. Fantasia[3], Kurt A. Püntener[3], R. David Britt[2,4], John F. Hartwig[1]*

Copper complexes are widely used in the synthesis of fine chemicals and materials to catalyze couplings of heteroatom nucleophiles with aryl halides. We show that cross-couplings catalyzed by some of the most active catalysts occur by a mechanism not previously considered. Copper(II) [Cu(II)] complexes of oxalamide ligands catalyze Ullmann coupling to form the C–O bond in aryl ethers by concerted oxidative addition of an aryl halide to Cu(II) to form a high-valent species that is stabilized by radical character on the oxalamide ligand. This mechanism diverges from those involving Cu(I) and Cu(III) intermediates that have been posited for other Ullmann-type couplings. The stability of the Cu(II) state leads to high turnover numbers, >1000 for the coupling of phenoxide with aryl chloride electrophiles, as well as an ability to run the reactions in air.

Metal-catalyzed cross-coupling reactions between aryl halide electrophiles and heteroatom nucleophiles are among the most widely used reactions for the discovery and production of fine chemicals and materials (*1*, *2*). Although copper catalysts were first used for such couplings more than a century ago (*3*, *4*), palladium catalysts led to the first mild couplings over 80 years later (*5*). Only recently, with the advent of new ligands, have couplings with catalysts based on the less expensive, earth-abundant metal copper become competitive with the counterparts catalyzed by palladium (*6*). However, much less is known about the mechanism of copper-catalyzed couplings than of palladium-catalyzed couplings, in part because copper complexes are often paramagnetic and, therefore, more difficult to identify spectroscopically; in addition, ligand-exchange processes are typically faster than the elementary steps of the catalytic cycle (*7*). Thus, most copper catalysts have been identified by trial and error.

Both the oxidation state of the active catalyst and the identity of the elementary steps on the catalytic cycle have been the source of much speculation (*8*). After initial mechanistic proposals that included σ-bond metathesis, activation of the carbon-halogen bond by π-coordination of the arene, or electron transfer to form aryl radicals (Fig. 1A) (*8*), careful studies have supported mechanisms in which a Cu(I) complex bound by commonly used neutral (*9*–*12*) or anionic (*12*, *13*) bidentate ligands undergoes concerted oxidative addition of an aryl halide (Fig. 1B). However, a variety of ligands proposed to bind as dianions also have been reported recently to generate catalysts for cross-coupling reactions that are challenging to achieve (Fig. 1C) (*14*–*16*). In particular, catalysts containing diamides derived from oxalic acid (oxalamide ligands), reported by Dawei Ma (*17*-*23*), are more active than any previous copper catalysts and enable reactions of unactivated, electron-rich aryl bromides and of aryl chlorides (*6*), but no mechanistic explanation of the high activity of the catalyst has been reported.

We show that complexes of oxalamide ligands catalyze the cross-coupling of aryl halides with phenols, and likely other nucleophiles, by a mechanism that is distinct from prior mechanisms considered for copper-catalyzed coupling. Our data imply that the mechanism involves a catalytically active, Cu(II) resting state and that a Cu(II) complex undergoes a concerted, rate-determining insertion of copper into the carbon-halogen bond of the aryl halide (Fig. 1D). Computations imply that radical character on the oxalamide ligand accommodates what would be an abnormally high valence at copper from a typical oxidative addition. The recognition of Cu(II) as the resting state led us to synthesize a preformed Cu(II) oxalamide complex, and this complex catalyzes couplings under air and coupling of an unactivated aryl chloride with turnover numbers exceeding 1000.

Synthesis of Cu(I) phenoxide complexes and reactions with aryl halides

Our mechanistic studies focused on reactions between aryl bromides and phenols to form biaryl ethers (*24*). The model coupling reaction of 1-bromo-4-fluorobenzene (**1**) with potassium phenoxide (**2**) formed 4-fluorobiphenyl ether (**3**) catalyzed by copper(I) iodide and bis-(2-phenylphenyl)oxalamide (H_2-BPPO, **4**) in 98% yield after heating at 100°C for 16 hours in dimethyl sulfoxide (DMSO) (Fig. 2A). Because studies with monoanionic ligands showed that couplings of phenols occur by turnover-limiting oxidative addition of aryl halides to a Cu(I) phenoxide complex (*13*), we first sought to generate a Cu(I) phenoxide complex bound by a mono- or dianionic form of the oxalamide ligand and to test the reactivity of such complexes with aryl halides.

Cu(I) complex **5** containing a dianionic oxalamide ligand was prepared from a 1:1 ratio of K,H-BPPO (**6**) and mesitylcopper(I). Single-crystal x-ray diffraction showed that complex **5** adopts a dimeric structure with two dianionic oxalamides bridging two copper atoms (supplementary materials section S2.4). The propensity of complex **5** to form a phenoxycopper complex (**8**) was assessed by progressive addition of potassium 3-*tert*-butylphenoxide to **5** in DMSO. Diagnostic ^{1}H nuclear magnetic resonance (NMR) signals of **5** shifted monotonically downfield, implying that phenoxide binds rapidly and reversibly to this complex.

The reaction of 1-bromo-4-fluorobenzene (**1**) with the mixture of **5** and potassium phenoxide at room temperature formed little aryl ether (Fig. 2B). Instead, the ^{1}H NMR resonances corresponding to the Cu complex **5** disappeared, along with 25% of **1**. No new ^{19}F NMR signals were observed, implying that the arene was transformed into a paramagnetic or insoluble species. The absence of ligand resonances suggests that the Cu(I) complex was oxidized to a new, paramagnetic, Cu(II) complex, and this hypothesis was confirmed by electron paramagnetic resonance (EPR) spectroscopy (supplementary materials section S4.7). Heating the sample at 100°C for 20 min formed biaryl ether **3** in 75% yield, along with fluorobenzene (**11**) in 18% yield, even though a Cu(II) species, rather than a Cu(I) phenoxide complex, was present in solution.

To determine whether an alternative cuprate **12** that bears one monoanionic oxalamide ligand, rather than a dianionic oxalamide, and one phenoxide ligand, was a catalytic intermediate, we prepared **12** by combining **5** with one equiv of 4-fluorophenol (**13**) (Fig. 2C). NMR spectroscopy and mass spectrometry suggest that heteroleptic cuprate **12** forms but equilibrates with the two homoleptic cuprates **14** and **15**. Addition of bromoarene **1** to this equilibrating mixture, again, caused the loss of ^{1}H NMR resonances of the ligand, presumably due to oxidation of Cu(I) to Cu(II). In this case, ^{19}F NMR spectroscopy showed that 0.20 equiv of **1** were consumed, and 0.10 equiv of fluorobenzene (**11**) formed. After heating at 100°C for 23 hours, 0.30 equiv of fluorobenzene (**11**) had formed, along with 0.20 equiv of biphenyl ether **16**; 0.33 equiv of **1** remained. Thus, the product distributions from the reaction of phenoxycuprates in the +1 oxidation state are far from the >95% yield from the

[1]College of Chemistry, University of California, Berkeley, Berkeley, CA 94720, USA. [2]Department of Chemistry, University of California, Davis, Davis, CA 95616, USA. [3]Pharmaceutical Division, Synthetic Molecules Technical Development, Process Chemistry and Catalysis, F. Hoffmann–La Roche, Ltd., Basel, CH-4070, Switzerland. [4]Miller Institute for Basic Research in Science, University of California, Berkeley, CA 94720, USA.
*Corresponding author. Email: jhartwig@berkeley.edu

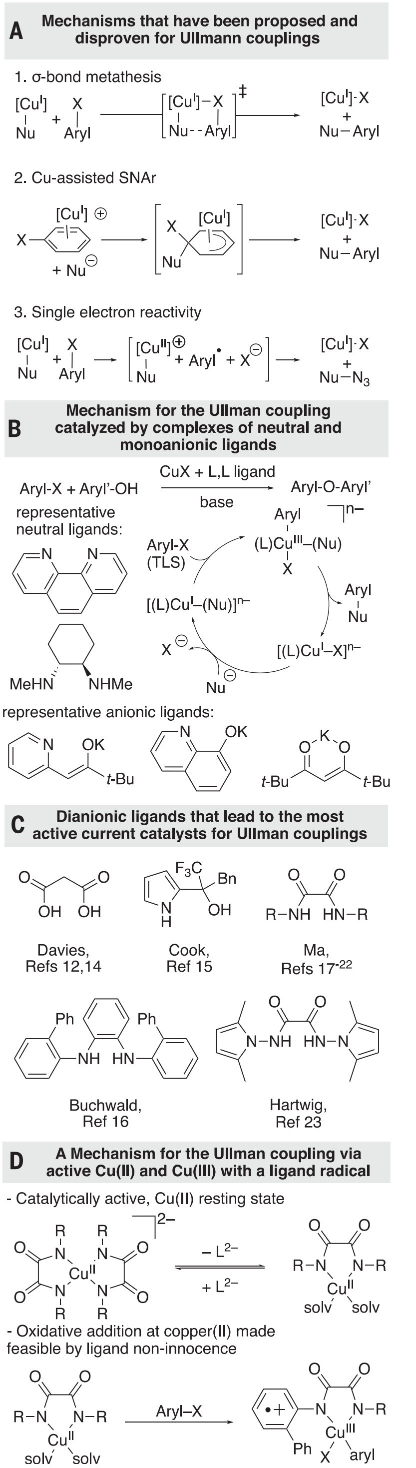
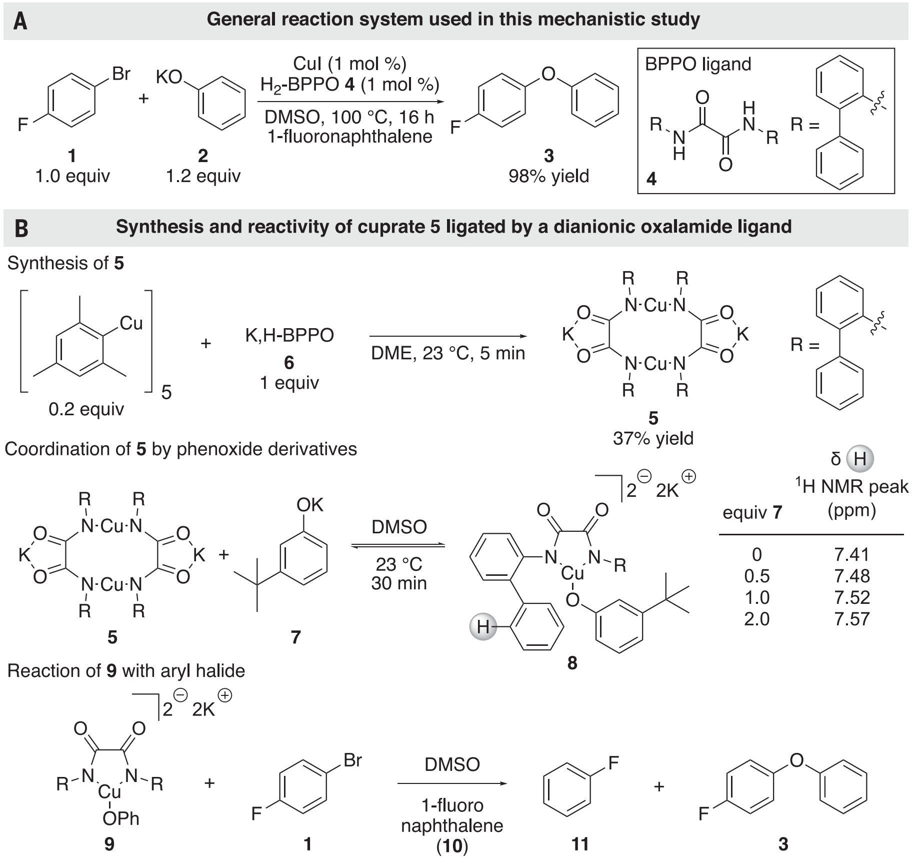
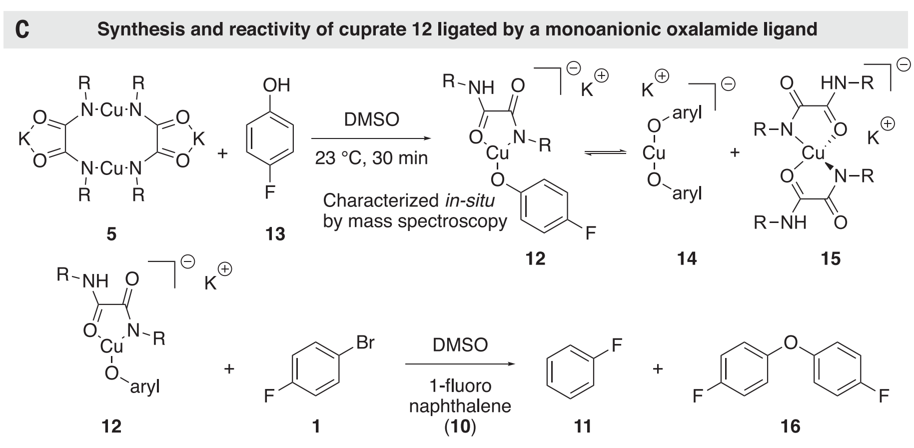


Fig. 1. Current understanding of reaction mechanisms for Cu-catalyzed C–O coupling reactions. (**A**) Proposed mechanisms for the reaction between copper complexes and aryl halides that have been ruled out by prior studies. (**B**) General reaction mechanism determined for C–O cross-coupling catalyzed by copper complexes of neutral or anionic bidentate ligands. (**C**) Recently discovered ligands that are highly active and can bind as dianionic ligands. (**D**) This work: Discovery of a mechanism with a copper(II) resting state and oxidative addition to a copper(II) complex.



time (min)	temp (°C)	equiv **9**	equiv **1**	equiv **11**	equiv **3**
0	23	1.00	0.00	0.00	0.00
15	23	0.00	0.75	0.00	0.00
20	100	0.00	0.00	0.18	0.75



time (min)	temp (°C)	equiv **12**	equiv **1**	equiv **11**	equiv **16**
0	23	0.80	1.00	0.00	0.00
15	23	0	0.81	0.10	0.00
20	100	0	0.60	0.20	0.12
1380	100	0	0.33	0.30	0.20

Fig. 2. Synthesis and reactivity of Cu(I) complexes of oxalamide ligands. (**A**) General conditions for the reaction under study, adapted from (*21*). (**B**) Synthesis of dianion-bound cuprate **5** and reactivity with aryl halides and potassium phenoxide derivatives. (**C**) Synthesis of monoanion-bound cuprate **12** and reactivity with aryl halide electrophiles.

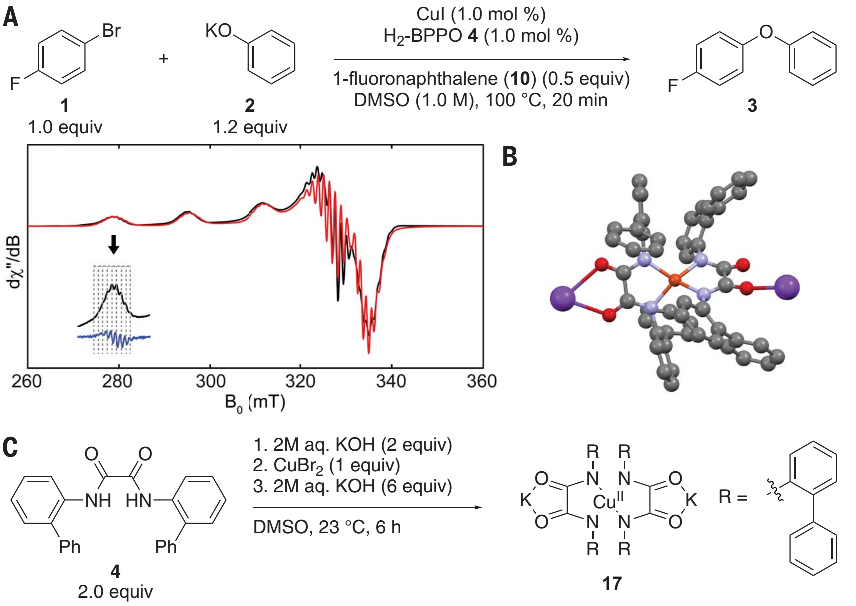


Fig. 3. Synthesis and key characterization of Cu(II) complex 17. (**A**) X-band EPR spectrum of a reaction mixture set up with CuI, H_2-BPPO, 4-fluorobromobenzene, potassium phenoxide, and 1-fluoronapthalene in DMSO, heated at 100°C for 30 min then flash frozen in liquid nitrogen. Black trace: Experimentally obtained spectrum. Red trace: Simulated spectrum. (Inset) The hyperfine splitting pattern within the $g_{\parallel}$ signal, along with the second-derivative spectrum in that region. (**B**) Structure of **17** determined by single-crystal x-ray diffraction; five DMSO molecules that cocrystallized and crystallographic disorder were removed for clarity. (**C**) Synthetic route to **17**.

catalytic coupling of the phenoxide with the same aryl halide, indicating that these Cu(I) species are not intermediates in the catalytic process.

Identification and reactions of Cu(II) complexes to assess their intermediacy in the catalytic reactions

To probe for Cu(II) complexes in the catalytic reaction, we obtained EPR spectra (Fig. 3 A). The EPR signals of a frozen reaction mixture generated from heating the aryl halide **1**, phenoxide **2**, copper iodide, and oxalamide **4** in DMSO at 100°C for 20 min corresponded to one Cu(II) species. The hyperfine pattern within the $g_{\parallel}$ signal at g = 2.210 (Fig. 3A, inset) indicates that the electron is coupled to four magnetically equivalent I = 1 nuclei. Copper(II) dioxalamide complex **17** bound by four nitrogen atoms is the most reasonable structure that matches the EPR data (*25*–*28*). Complex **17** was independently synthesized from $CuBr_2$, ligand, and base (Fig. 3C) and was fully characterized, including single-crystal x-ray diffraction (Fig. 3B). The EPR spectrum of this material matched that of the catalytic reaction (supplementary materials section S4.1 to S4.5).

Because a Cu(II) complex can initiate reactions catalyzed by a Cu(I)/Cu(III) couple (*29*), we assessed whether the Cu(II) complex **17** is a true catalyst or a precatalyst that is reduced to Cu(I) (Fig. 4A). To do so, complex **17** and one of the two organic reactants were combined in an equimolar ratio, along with a 10-fold excess of the coupling partner. If complex **17** were a precatalyst and transforms to the active catalyst more rapidly than the coupling reaction, then the reagent in equimolar amounts with copper would be consumed in the catalyst activation step, leading to a low yield of biaryl ether. If complex **17** transforms to the active catalyst more slowly than the coupling reaction, then an induction period would be observed. However, if the observed Cu(II) species were a true intermediate, then the reaction would form a nearly quantitative yield of biaryl ether without an induction period and with a reaction progress curve that would reflect the order of the limiting reagent.

In the event, the reaction of Cu(II) **17** with one equiv of phenoxide **18** and excess of aryl halide **1** gave a 97% yield of biaryl ether **16** with a pseudo–zero order reaction profile through 70% conversion. The reaction of Cu(II) **17** with one equiv of aryl halide **1** and excess of phenoxide **18** gave a 99% yield of biaryl ether **16** with a pseudo–first order reaction profile. These results are inconsistent with the conversion of the Cu(II) complex to an active Cu(I) species and are consistent with the intermediacy of **17** or a species formed reversibly from **17**.

Kinetic studies were performed to establish the relative stoichiometry of the resting state and the turnover-limiting transition state (Fig. 4B). The reaction was zero order in potassium 4-chlorophenoxide (**19**), indicating that the phenoxide is either present in the resting state or that it enters the catalytic cycle after the turnover-limiting step. The reaction was nearly first order in aryl halide **1**, indicating that it associates prior to the transition state of the catalytic process. Added dianionic ligand K_2-BPPO (**20**) inhibited the reaction but had no measurable effect on the rate at concentrations higher than 0.50 mM. Because added **20** does not coordinate to **17** in solution (supplementary materials section S4.6), these data are consistent with a combination of reaction pathways involving dissociative and associative displacement of a ligand by the aryl halide. Finally, the reaction occurred with a linear dependence on the concentration of catalyst, albeit with an apparent nonzero intercept. These kinetic data are discussed in further detail in the supplementary materials.

Our experiments rule out reduction of **17** by phenoxide to form Cu(I), but we also conducted cyclic voltammetry (CV) experiments to address whether two equivalents of Cu(II) could disproportionate to form a catalytically active Cu(I) species and a corresponding Cu(III) species. CV was conducted in DMSO containing 0.20 M tetrabutylammonium hexafluorophosphate as electrolyte (Fig. 4C). The oxidation of **17** to a Cu(III) species and the corresponding reduction of **17** to a Cu(I) species are both reversible and separated by 1.81 V. This separation shows that disproportionation of two equiv of **17** into a Cu(I) and Cu(III) species is uphill by more than 41 kcal/mol, providing strong evidence against the generation of the active catalyst by disproportionation of the Cu(II) species.

Distinguishing between four potential mechanisms

Four mechanisms that have been proposed previously for Ullmann couplings and that are consistent with the kinetic data are shown in Fig. 5A. These experiments include a sigma-bond metathesis process by which a nucleophile-bound copper complex reacts directly with the aryl halide in a redox-neutral step; a nucleophilic aromatic substitution reaction that is also redox neutral and is facilitated by formation of an arene complex to Cu; electron transfer from Cu to the aryl halide to form a Cu(III)

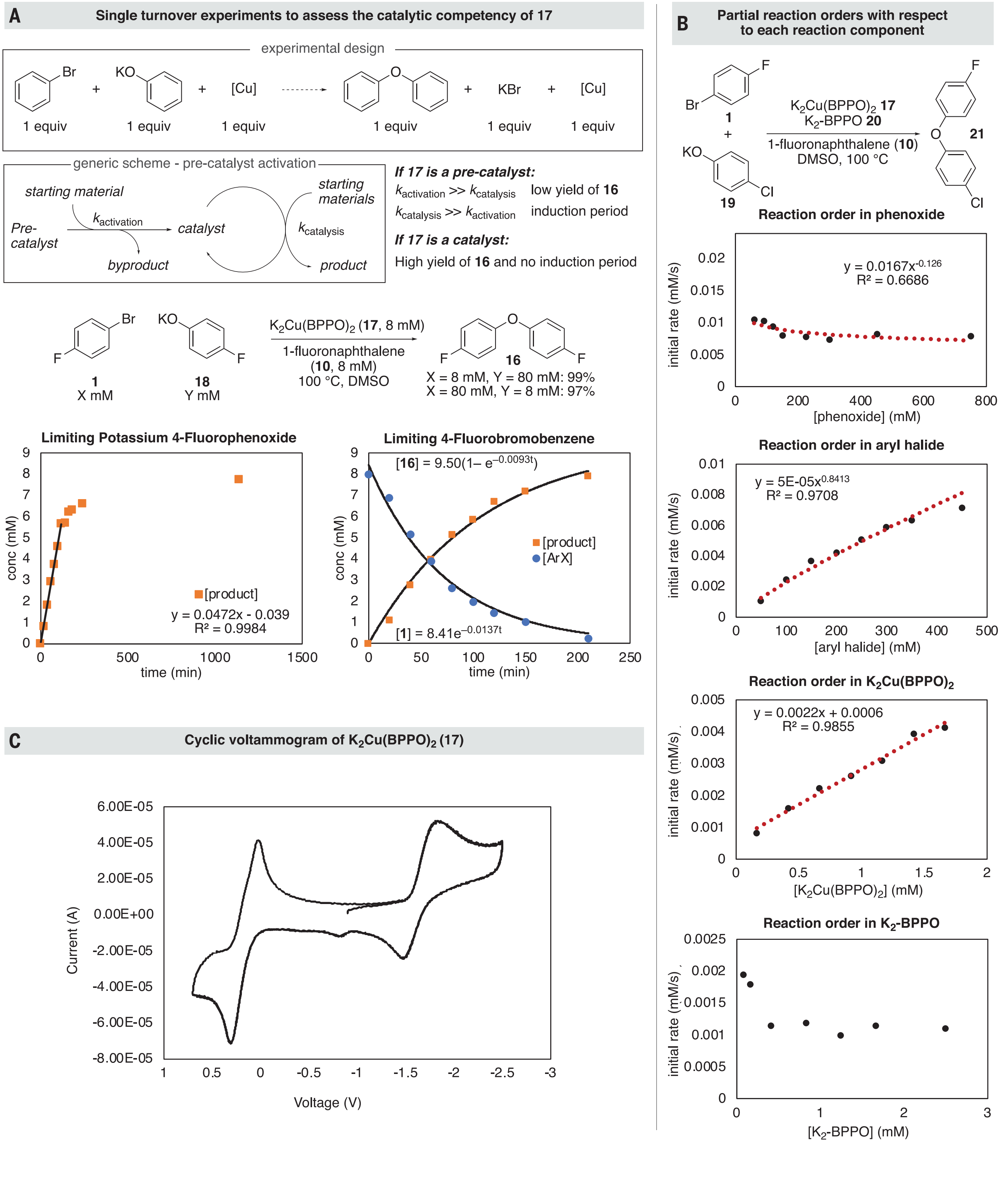


Fig. 4. Mechanistic experiments to establish the resting state, turnover-limiting step, and pre-equilibria. (**A**) Single-turnover experiment that determines whether **17** must be activated by either coupling partner before entering the catalytic cycle. (**B**) Reaction orders in phenoxide (zero order), aryl halide (first order), catalyst (first order), and ligand dipotassium salt (inverse order). (**C**) Cyclic voltammogram of $K_2Cu(BPPO)_2$ in a 0.20 M DMSO solution of tetrabutylammonium hexafluorophosphate.

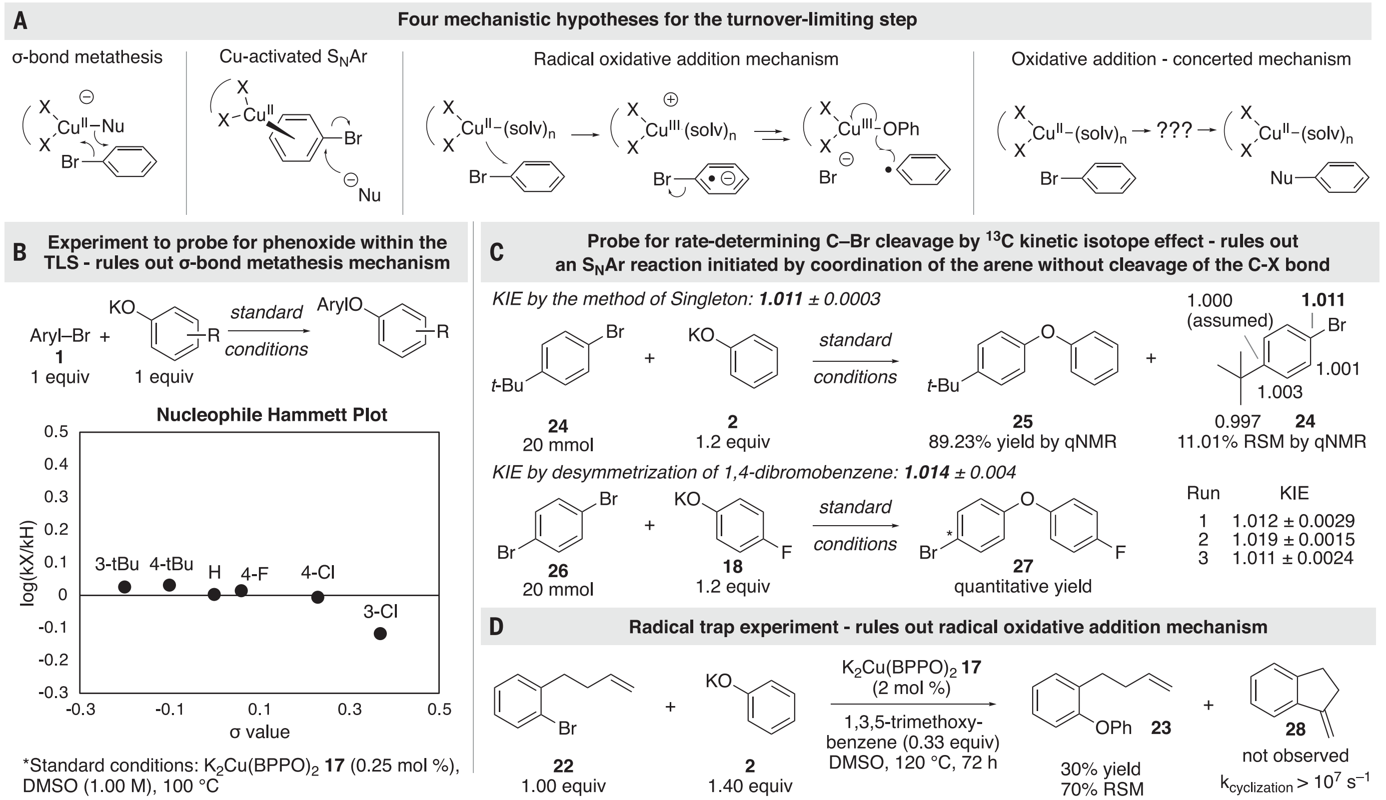


Fig. 5. Mechanistic studies to probe the nature of the resting state and turnover-limiting step. (**A**) Four mechanistic hypotheses for the turnover-limiting reaction between Cu and the aryl halide. (**B**) Hammett plot demonstrating that phenoxide is not bound to copper in the turnover-limiting step. (**C**) Natural abundance ^{13}C kinetic isotope experiments conducted by the methods of Singleton and Vetticatt. (**D**) Reactions conducted with an aryl bromide containing a pendant alkene that probes the intermediacy of aryl radical species.

intermediate and an aryl radical, which combines with bound phenoxide; and oxidative addition by a concerted pathway resulting in a high-valent copper intermediate. A series of further experiments were conducted to distinguish these mechanisms.

The sigma-bond metathesis mechanism would be consistent with our kinetic data only if the phenoxide were bound to copper in the resting state and remained bound through the turnover-limiting step. To assess whether a copper phenoxide complex reacts with the aryl halide, we measured the rate as a function of the identity of the phenoxide. If phenoxide were bound when the copper species reacts with the aryl halide, then the electronic properties of the phenoxide would change the identity of the complex and influence the rate of the reaction. However, no influence of the identity of the phenoxide on the rate of the reaction was detected (Fig. 5B). Further studies that show an effect of phenoxide electronic properties on the steps after the rate-determining step, thereby corroborating a lack of binding of phenoxide derivatives to Cu(II) complex **17**, are described in the supplementary materials.

The S_NAr pathway involving rate-determining attack of phenoxide on a bound aryl halide is inconsistent with the zeroth reaction order in phenoxide, but a version of this mechanism involving rate-limiting coordination of arene before nucleophilic attack is consistent with the data presented so far. To determine if coordination of the arene or cleavage of the C-X bond occurs during the rate-limiting reaction between Cu(II) and the aryl halide, we measured the ^{13}C kinetic isotope effect (KIE) (Fig. 5C). No ^{13}C KIE at the ipso carbon would be expected for reaction by the mechanism involving irreversible coordination of the aryl halide, whereas a primary ^{13}C KIE at this carbon would be expected for a mechanism involving irreversible cleavage of the aryl halide by copper. ^{13}C KIE values were measured by both intermolecular competition (*30*) and intramolecular competition (*31*) (1.011 ± 0.0003 and 1.014 ± 0.004, respectively). These normal, primary ^{13}C KIE values are comparable to those observed previously for oxidative additions of aryl iodides to Cu(I) complexes (*13*) and rule out irreversible binding of the aryl halide prior to nucleophilic aromatic substitution by the phenoxide.

The third pathway involving electron transfer to form an aryl halide radical anion that generates halide and an aryl radical was assessed by conducting the coupling of phenoxide with aryl halide **22** catalyzed by Cu(II) complex **17** (Fig. 5D). An aryl radical formed from **22** is known to cyclize rapidly ($k > 10^7\ s^{-1}$) (*32*) with the *ortho*-butenyl substituent. However, the reaction of phenoxide **2** with aryl halide **22** formed no cyclized products, as determined by gas chromatography–mass spectrometry and NMR spectroscopy. This result, in concert with the low reducing power of complex **17** (Fig. 4C), rules out reduction of the aryl halide to generate an aryl radical intermediate and subsequent outer-sphere reductive elimination with a copper(III) phenoxide complex, and it is consistent with the previously reported lack of cyclization during the reaction between 2-butenyl-chlorobenzene and *p*-cresol catalyzed by copper and a different oxalamide ligand (*18*). Instead, all our data are consistent with the fourth mechanism involving concerted, rate-determining reaction of the aryl halide with a Cu(II) complex of the oxalamide ligand to cleave the carbon-halogen bond.

DFT study of the oxidative addition of aryl halides to Cu(II).

We assessed by density functional theory (DFT) calculations the mechanism by which a Cu(II) species could react with an aryl halide to

cleave the carbon-halogen bond. Because large differences in energy result from modeling of charged species with different degrees of explicit solvation, we did not compute the energetics for dissociation of the anionic ligand from the cuprate to form a neutral Cu(II) species. Instead, we studied the energetics for binding of the aryl bromide to (BPPO)Cu (**29**) and the feasibility of an insertion of the Cu intermediate into the carbon-halogen bond of the bound aryl halide by a step akin to a concerted oxidative addition. Analysis by the fractional orbital density (FOD) method published by Grimme (*33*) demonstrated that most structures on the reaction coordinate possess appreciable multireference character (N_FOD = 1.4 to 2.0). Because the inclusion of HF-exchange energy causes high errors in the relative calculated energies for structures with multireference character (*33*), we used the meta-GGA functional TPSS-D3(BJ) for all calculations. Geometry optimization and frequency calculations were conducted with the def2-TZVP basis set on copper and the def2-SVP basis set on all other atoms, and single-point energies were calculated with the def2-TZVP basis set on all atoms. The computed profile begins with a copper(II) complex ligated by an *N*,*N*-bound oxalamide ligand (**29**) (Fig. 6A). Association of the aryl halide to yield **30** was calculated to be endergonic by 12.4 kcal/mol, and the product from oxidative addition of the bound arene (**31**) was found to be 21.0kcal/mol higher in free energy than the starting complex. The transition state to form this complex (**32**) was computed to lie 23.1 kcal/mol above the starting complex. These energies are consistent with our conclusion from experimental studies that the Cu(II) species cleaves the carbon-halogen bond in the rate-determining step and suggest that this bond cleavage can occur by oxidative addition. A full account of our DFT studies is included in the supplementary materials.

To reveal the electronic structure of the complexes involved in this oxidative addition, we conducted natural population analysis on each species. Arylcopper halide complex **31** lacked spin density on copper. Instead, >98% of the spin density was located on the ligands bound to copper, with 86% located on the oxalamide. This ligand redox non-innocence (*34*) enables the Cu(II) resting state to undergo oxidative addition of the aryl halide to form complex **31** containing both an aryl ligand and a halide ligand without a nearly unprecedented oxidation state at Cu. Indeed, natural population analysis of this complex suggests that the copper in it possesses 10 d-electrons, and this electron count implies that intermediate **31**, although formally Cu(III) by the presence of three formally anionic ligands and an overall neutral charge, is best described as a Cu(I) complex with an inverted ligand field, as discussed for many formally Cu(III) complexes (*35*, *36*), and with radical character on the ligand to accommodate the large number of X-type ligands.

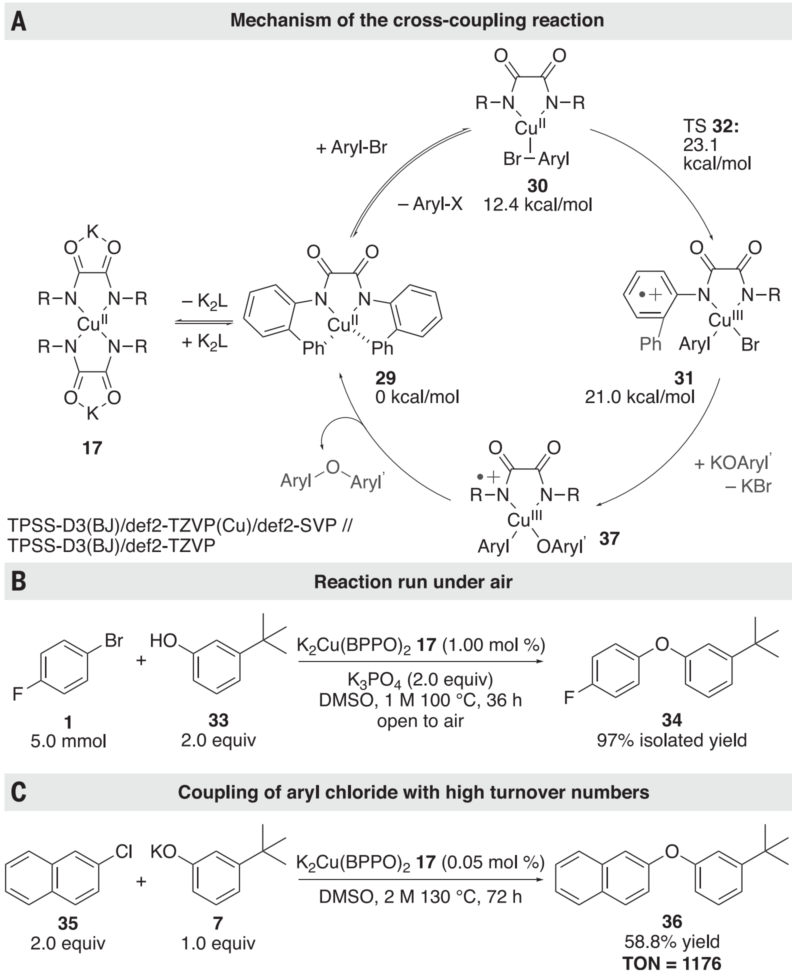


Fig. 6. Mechanism of the cross-coupling reaction and synthetic applications. (**A**) Catalytic cycle representing our mechanistic conclusions. (**B** and **C**) Applications of the Cu(II) complex in preparative, cross-coupling reactions.

Reaction improvements with a preformed Cu(II) catalyst

A Cu(II) resting state is stable to air and disproportionation. Thus, a purposeful synthesis of a Cu(II) catalyst should create a system that is stable to air and that reacts with high turnover numbers. Indeed, the cross-coupling reaction between 4-fluorobromobenzene (**1**) and 3-*tert*-butyl phenol (**33**) catalyzed by just 1.0 mol % of **17** in DMSO on a 1-mmol scale without any effort to exclude air or water gave the coupled product **34** in 97% yield after 36 hours at 100°C. Moreover, the coupling of 2-chloronaphthalene (**35**) with potassium 3-*tert*-butylphenoxide (**7**) with just 500 parts per million of pure **17** as the catalyst for 72 hours at 130°C gave 59% yield of biaryl ether **36**, corresponding to >1000 turnovers. We propose that the high stability of these catalysts at the temperatures required for the activation of carbon-chlorine bonds makes them more long-lived for this transformation than electron-rich, Cu(I) complexes that are more prone to degradation.

These results highlight the potential of well-defined copper(II) catalysts to increase catalyst activity, turnover numbers, and reproducibility of copper-catalyzed coupling reactions with this and related classes of ligands. In addition, the discovery of this reaction mechanism could facilitate the development of next-generation ligands for catalysts containing copper or other first-row metals to minimize single-electron reactivity and allow such metals to fill roles that are currently realized by noble metal catalysts in a variety of reactions important for organic synthesis.

REFERENCES AND NOTES

1. R. Frian, D. Kikelj, *ChemInform* **37**, 200641247 (2006).
2. P. Ruiz-Castillo, S. L. Buchwald, *Chem. Rev.* **116**, 12564–12649 (2016).

3. F. Ullmann, P. Sponagel, *Ber. Dtsch. Chem. Ges.* **38**, 2211–2212 (1905).
4. I. Goldberg, *Ber. Dtsch. Chem. Ges.* **39**, 1691–1692 (1906).
5. J. F. Hartwig, K. H. Shaughnessy, S. Shekhar, R. Green, *Org. React.* **100**, 14 (2019).
6. S. Bhunia, G. G. Pawar, S. V. Kumar, Y. Jiang, D. Ma, *Angew. Chem. Int. Ed.* **56**, 16136–16179 (2017).
7. J. E. Hein, V. V. Fokin, *Chem. Soc. Rev.* **39**, 1302–1315 (2010).
8. C. Sambiagio, S. P. Marsden, A. J. Blacker, P. C. McGowan, *Chem. Soc. Rev.* **43**, 3525–3550 (2014).
9. J. W. Tye, Z. Weng, R. Giri, J. F. Hartwig, *Angew. Chem. Int. Ed.* **49**, 2185–2189 (2010).
10. J. W. Tye, Z. Weng, A. M. Johns, C. D. Incarvito, J. F. Hartwig, *J. Am. Chem. Soc.* **130**, 9971–9983 (2008).
11. R. Giri, J. F. Hartwig, *J. Am. Chem. Soc.* **132**, 15860–15863 (2010).
12. Q. A. Lo, D. Sale, D. C. Braddock, R. P. Davies, *ACS Catal.* **8**, 101–109 (2018).
13. R. Giri *et al*., *J. Am. Chem. Soc.* **140**, 793–806 (2018).
14. S. Sung *et al*., *ACS Catal.* **6**, 3965–3974 (2016).
15. A. de Gombert *et al*., *ACS Catal.* **13**, 2904–2915 (2023).
16. S.-T. Kim, M. J. Strauss, A. Cabré, S. L. Buchwald, *J. Am. Chem. Soc.* **145**, 6966–6975 (2023).
17. W. Zhou, M. Fan, J. Yin, Y. Jiang, D. Ma, *J. Am. Chem. Soc.* **137**, 11942–11945 (2015).
18. M. Fan, W. Zhou, Y. Jiang, D. Ma, *Angew. Chem. Int. Ed.* **55**, 6211–6215 (2016).
19. S. De, J. Yin, D. Ma, *Org. Lett.* **19**, 4864–4867 (2017).
20. Z. Chen, D. Ma, *Org. Lett.* **21**, 6874–6878 (2019).
21. J. Gao *et al*., *Org. Lett.* **19**, 2809–2812 (2017).
22. Y. Chen, L. Xu, Y. Jiang, D. Ma, *Angew. Chem. Int. Ed.* **60**, 7082–7086 (2021).
23. R. Ray, J. F. Hartwig, *Angew. Chem. Int. Ed.* **60**, 8203–8211 (2021).
24. Y. Zhai *et al*., *J. Org. Chem.* **82**, 4964–4969 (2017).
25. For prior literature on Cu(II) complexes of oxalamide ligands, see (*26*–*28*).
26. H. Ojima, K. Nonoyama, *Coord. Chem. Rev.* **92**, 85–111 (1988).
27. H. Desseyn, G. Schoeters, *Bull. Soc. Chim. Belg.* **95**, 13–27 (1986).
28. B. Cervera *et al*., *J. Chem. Soc., Dalton Trans.*781–790 (1998).
29. A. J. Paine, *J. Am. Chem. Soc.* **109**, 1496–1502 (1987).
30. D. A. Singleton, A. A. Thomas, *J. Am. Chem. Soc.* **117**, 9357–9358 (1995).
31. V. Wambua, J. S. Hirschi, M. J. Vetticatt, *ACS Catal.* **11**, 60–67 (2021).
32. M. D. Koppang, G. A. Ross, N. F. Woolsey, D. E. Bartak, *J. Am. Chem. Soc.* **108**, 1441–1447 (1986).
33. S. Grimme, A. Hansen, *Angew. Chem. Int. Ed.* **54**, 12308–12313 (2015).
34. V. Lyaskovskyy, B. de Bruin, *ACS Catal.* **2**, 270–279 (2012).
35. I. M. DiMucci *et al*., *J. Am. Chem. Soc.* **141**, 18508–18520 (2019).
36. R. Hoffmann *et al*., *Chem. Rev.* **116**, 8173–8192 (2016).

ACKNOWLEDGMENTS

We thank E. Kalkman, A. Zahrt, D. Hait, and J. Brunn for helpful discussions regarding density functional theory; R. Chatterjee and C. Asokan for assistance with EPR spectroscopy experiments; and Z. Zhao and A. Wheeler for assistance collecting high-resolution mass spectrometry data of Cu(I) complexes. We thank H. Celik and UC Berkeley's NMR facility in the College of Chemistry (CoC-NMR) for spectroscopic assistance. Instruments in the CoC-NMR are supported in part by NIH S10OD024998. We thank K. Durkin and D. Small at UC Berkeley's Molecular Computation and Graphics Facility (MGCF). Computing resources at the MGCF are supported in part by NIH S10OD034382. We thank N. Settineri and Berkeley's College of Chemistry X-ray Crystallography Facility (CheXray) for assistance with crystallography. CheXray is supported in part by NIH S10RR027172. **Funding:** Funding for this work was provided in part by F. Hoffmann-La Roche Ltd, in part through NIH F32GM140550, and in part through NIH 1R35GM126961-01. C.P.D. is supported by NIH Kirschstein NRSA postdoctoral fellowship F32GM140550. I.F.Y. is supported by an NSF GRFP fellowship under DGE 1752814 and DGE 2146752. **Author contributions:** C.P.D., Conceptualization (supporting), investigation (lead), methodology (lead), original draft preparation (lead), review and editing (equal); E.L., Investigation (supporting, chemical synthesis and mechanistic experiments); Q.H., Investigation (supporting, DFT); I.F.Y., Investigation (supporting, x-ray crystallography); L.T., Investigation (supporting, EPR spectroscopy); G.R., Investigation (supporting, EPR spectroscopy); A.J., Investigation (supporting, chemical synthesis and mechanistic experiments); S.M.F., Supervision (supporting), review and editing (equal); K.A.P., Supervision (supporting); R.D.B., Supervision (supporting); J.F.H., Conceptualization (lead), supervision (lead), funding acquisition (lead), review and editing (equal). **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All other data are available in the main text or the supplementary materials. X-ray data are available in the Cambridge Crystallographic Data Center under CCDC numbers 2265761 and 2265762. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adi9226
Materials and Methods
Supplementary Text
Figs. S1 to S36
Tables S1 to S36
NMR Spectra
References (*37*–*57*)

Submitted 27 May 2023; accepted 8 August 2023
10.1126/science.adi9226

VOLCANOLOGY

Fast and destructive density currents created by ocean-entering volcanic eruptions

Michael A. Clare[1]*†, Isobel A. Yeo[1]*†, Sally Watson[2], Richard Wysoczanski[2], Sarah Seabrook[2], Kevin Mackay[2], James E. Hunt[1], Emily Lane[2], Peter J. Talling[3], Edward Pope[3], Shane Cronin[4], Marta Ribó[5], Taaniela Kula[6], David Tappin[7], Stuart Henrys[8], Cornel de Ronde[8], Morelia Urlaub[9], Stefan Kutterolf[9], Samuiela Fonua[10], Semisi Panuve[10], Dean Veverka[11], Ronald Rapp[12], Valey Kamalov[13], Michael Williams[2]

Volcanic eruptions on land create hot and fast pyroclastic density currents, triggering tsunamis or surges that travel over water where they reach the ocean. However, no field study has documented what happens when large volumes of erupted volcanic material are instead delivered directly into the ocean. We show how the rapid emplacement of large volumes of erupted material onto steep submerged slopes triggered extremely fast (122 kilometers per hour) and long-runout (>100 kilometers) seafloor currents. These density currents were faster than those triggered by earthquakes, floods, or storms, and they broke seafloor cables, cutting off a nation from the rest of the world. The deep scours excavated by these currents are similar to those around many submerged volcanoes, providing evidence of large eruptions at other sites worldwide.

Explosive volcanism poses a wide range of hazards, with more than a third of volcanic fatalities attributed to fast (up to hundreds of kilometers per hour) and high-temperature pyroclastic density currents triggered by phreatic explosions, pyroclastic fountaining, lateral blasts, caldera and dome collapses, and the vertical collapse of eruption columns where they contact the ground (*1*–*6*). Study of the behavior of pyroclastic density currents on land has revealed a spectrum from dense to dilute and turbulent modes of flow, across which a range of hazards exists (*1*–*3*). This spectrum also relates to other types of particulate density currents, including snow avalanches and underwater sediment–laden flows called turbidity currents (*7*–*9*). Where terrestrially initiated pyroclastic density currents reach the ocean, they create different hazards. Such currents can generate tsunamis, create phreatic explosions as hot currents interact with water, travel over the sea, and/or rapidly cool and transition into a turbidity current, damaging seafloor infrastructure and devastating marine biological communities (*10*–*15*).

Repeat terrestrial surveys and sampling after the occurence of pyroclastic density currents have enabled reconstruction of flow properties from the resultant deposits, characterizing the associated hazards (*16*–*18*). However, field-scale observations of density currents linked to volcanic eruptions in marine settings are rare, owing to often-remote locations and inaccessibility of in situ deposits. The behavior of terrestrially initiated pyroclastic density currents that cross land to enter the ocean has only been documented at a single field site after a small (0.19 km^3) eruption (*12*–*14*), whereas no equivalent study has shown what happens when an eruption directly delivers volcanic material into the ocean. We address this knowledge gap with observations of underwater volcaniclastic density currents triggered during the eruption of the partially submerged Hunga Tonga–Hunga Ha'apai volcano (hereafter referred to as Hunga volcano) in the Kingdom of Tonga. We use the term "volcaniclastic density current," which encompasses a spectrum of density currents linked to a volcanic eruption, from hot gas–supported pyroclastic density currents to cold fluid–supported turbidity currents.

A key control on the hazard posed by any type of density current is its velocity (*9*, *19*–*22*). Although recent advances in technology have enabled the direct measurement of turbidity currents and snow avalanches, no velocity measurements exist for underwater volcaniclastic density currents (*9*, *19*–*21*). These limitations are notable, considering the distinct hazards posed by partially or fully submerged volcanoes, which account for more than three-quarters of active volcanoes worldwide (*6*). Consequently, our knowledge relies on studies of ancient ocean-entering eruptions (*23*, *24*), scaled-down laboratory modeling (*25*), and analysis of geomorphic features around submerged volcanoes to infer the behavior of past eruptions (*26*, *27*). Fields of large sediment waves and scours, commonly observed radiating around submerged flanks of volcanoes, are thought to be diagnostic of catastrophic eruptions (*26*–*28*). However, this hypothesis remains untested because of a lack of repeat seafloor surveys before and after a large eruption. These uncertainties severely limit the understanding of the behavior and associated risks at submerged volcanoes.

We present observations of voluminous volcaniclastic density currents that were triggered by the 15 January 2022 eruption of Hunga volcano in the Kingdom of Tonga. This eruption was the most explosive in more than a century and had worldwide impacts (*29*–*35*). The eruption plume entered the mesosphere (57 km high), tsunamis traveled across the Pacific Ocean and caused 19- to 20-m runups in Tonga, and a pressure wave encircled the globe multiple times (*29*–*31*, *33*, *34*). More than 1 hour after the main eruption, Tonga's only international subsea telecommunications cable was severed (Fig. 1), disconnecting the entire nation from global digital communications at a critical time for disaster response (*36*). Such an incident has wider implications because subsea cables carry >99% of all international data traffic, including the internet and trillions of dollars per day in financial transactions (*37*). The >6 km^3 eruption expelled a volume of material equivalent to the annual sediment flux from all the world's rivers combined, much of which directly entered the ocean through eruption column collapses (*15*, *38*). The rapid escalation in explosivity and the resultant hazards were unexpected, exposing a gap in understanding of many similar, yet unmonitored, volcanoes along the Tonga-Tofua arc and volcanic settings worldwide (*6*, *29*, *30*, *39*).

By integrating datasets that document their timing and extent, we determined the behavior of underwater volcaniclastic density currents triggered by the 2022 eruption of Hunga volcano. These currents traveled >100 km and caused extensive damage to seafloor cables, from which we could estimate their velocity, which reached up to 122 km/hour. These currents differ markedly from those triggered by terrestrially initiated pyroclastic density currents that enter the ocean, and their velocity is higher than that recorded for other underwater density currents, including those triggered by terrestrial floods, large earthquakes, or ocean waves (Table 1). The currents created distinctive scours around Hunga volcano, excavating >100-m-deep regions into the volcanic edifice (Fig. 2), which were evident by comparing surveys 3 months after the eruption to pre-eruption surveys performed in 2016. Bedforms like these

[1]National Oceanography Centre, Southampton, UK. [2]National Institute of Water and Atmospheric Research (NIWA), Auckland, Aotearoa New Zealand. [3]Department of Geography and Department of Earth Sciences, Durham University, Durham, UK. [4]School of Environment, University of Auckland, Auckland, Aotearoa New Zealand. [5]Department of Environmental Science, Auckland University of Technology, Auckland, Aotearoa New Zealand. [6]Ministry of Lands and Natural Resources, Nuku'alofa, Kingdom of Tonga. [7]British Geological Survey, Keyworth, UK. [8]GNS Science, Lower Hutt, Aotearoa New Zealand. [9]GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany. [10]Tonga Cable Ltd, Nuku'alofa, Kingdom of Tonga. [11]Southern Cross Cable Network, North Ryde, New South Wales, Australia. [12]SubCom, Newington, NH, USA. [13]Valey Kamalov LLC, Gainesville, FL, USA.
*Corresponding author. Email: michael.clare@noc.ac.uk (M.A.C.); i.yeo@noc.ac.uk (I.A.Y.)
†These authors contributed equally to this work.

are generated by, and thus probably record, large explosive eruptions.

Extensive damage to seafloor cables

At 03:47 (all times UTC) on 15 January 2022, a low eruptive plume was observed at Hunga volcano, marking the onset of the main eruptive phase, with a steady narrow plume rising to a height of >10 km (*34*) (Fig. 1). At 04:15 (T_{0a}), a large explosion occurred [volcanic explosivity index of between 5 and 6 (*35*)], which transformed the plume. High eruption rates rapidly increased the height and width of an expanding umbrella-shaped eruption plume (*34*) (fig. S6). The plume reached a height of between 16 and 18 km at 04:17. A second major explosion (T_{0b}) occurred at 04:21, with additional explosions generating sonic booms until 04:25 (inception of the atmospheric pressure wave). By this time, the umbrella cloud had expanded to a width of at least 120 km, and the central plume was >15 km wide. Plume collapses into the ocean below the umbrella cloud occurred soon after 04:17, especially from 04:20 onward (*34*) (fig. S6). By 04:50 to 04:55, the central high plume ceased rising and was dispersed by the wind. However, the eruption continued vigorously with a lower plume (~17 to 21 km high) formed beneath the larger 57-km-high plume.

Both subsea cables laid near Hunga volcano were broken on 15 January, but the timing of this damage lagged after the two most intense explosions (T_{0a} and T_{0b}) by 9 to 15 min for the domestic cable (at 04:30) and by 83 to 89 min for the international cable (05:44) (table S1). The timing of these breaks is known to the nearest minute, on the basis of loss of data transmission, ultimately with complete loss of internet capacity when the international cable was severed. The distance of the first point of cable damage from shore was determined immediately using optical time-domain reflectometry, but the full extent could not be assessed until a cable repair ship retrieved the intact ends of the cable on either side of the damaged zone. The international cable repair took 5 weeks, as the closest repair ship was 2500 km away, in Papua New Guinea, and >100 km of replacement cable was required. Communications were limited across the kingdom until the domestic cable was finally repaired, 18 months after the eruption.

Prior to repair, cable damage was thought to be caused by local seabed landslides; however, the extent of damage was far greater than anticipated. More than 89 km of the international cable was broken and/or buried to a depth beyond recovery, while 105 km of the domestic cable was affected. Moreover, the international cable was recovered at a distance of 5 km to the north of its originally laid position, closer toward the volcano. This incident was the largest length

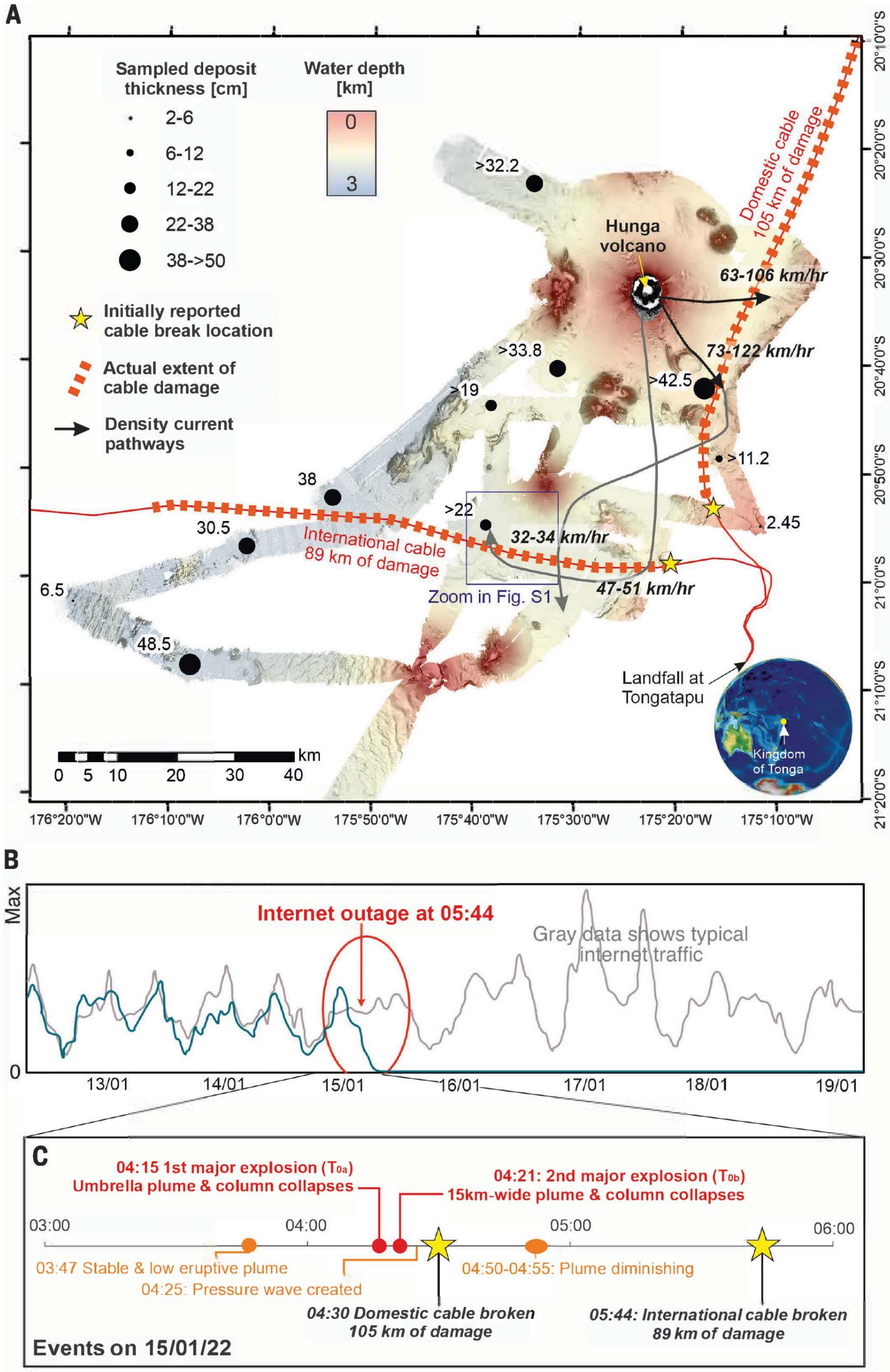


Fig. 1. Extensive damage caused to seafloor cables by volcaniclastic density currents generated by column collapse at Hunga volcano. (**A**) Locations and extent of cable damage on the domestic and international seafloor cables resulting from the volcaniclastic density currents (pathways shown as arrowed lines) plotted on bathymetric data acquired 3 months after the eruption. The thickness of volcaniclastic density current deposits as sampled by multicoring is depicted as size-scaled solid circles, showing that these currents deposited material at least 108 km away from the caldera. Actual sampled thickness of volcaniclastic density current deposit is annotated. Where the base of the density current deposit was not sampled, the thickness is given as >X cm. (**B**) Internet capacity shown for typical periods (in gray) compared with the sudden loss of internet traffic, which flatlined at 05:44 on 15 January 2022, when the international cable was broken. (**C**) Enhanced timeline of the main eruptive phase of Hunga volcano on 15 January 2022, including the two major eruptions that caused ocean-entering column collapses. The timings of the two cable breaks are marked with stars, showing that they occurred after the main explosive eruptions.

of cable damage since telegraph cables were buried by the >200-km^3 Grand Banks landslide offshore Newfoundland in 1929 (*40*) and the longest single cable repair in the modern fiber-optic era (Fig. 1). We did not find evidence within the resolution of imaging of slope failures on the deep-water volcano flanks around the cables and surrounding slopes. We therefore relate the cable damage to powerful and long runout volcaniclastic density currents triggered by the effective delivery of large volumes of erupted pyroclastic material directly into the ocean.

Insights into underwater density currents triggered by eruption column collapse

Dense and highly erosive proximal regime

Linear gullies and trains of crescentic scours, incised up to 100 m, radiate from the caldera (Fig. 2), accounting for 3.5 km^3 of erosion (*15*) and representing an additional removed mass of >50% of the original erupted volume. Erosional features include scours of up to 2 km in width and upslope-asymmetric bedforms (30 to 60 m wave height, 500 to 2000 m wavelength) seen in pre-eruption surveys, which migrated up to 1.5 km upslope. Pronounced erosion observed from the post-eruption survey occurs within a radius of <9.2 km from the caldera rim and only on the steepest slopes (>10°, locally >40°; Fig. 2). Recovery of material by seafloor coring was not successful on such proximal slopes, presumably owing to the presence of competent

Table 1. Reported velocities of fast-moving (>4 m/s) submerged particulate density currents worldwide. Values based on sequential seafloor cable breaks or acoustic monitoring. Also compared are subaerial density currents, including pyroclastic density currents and snow avalanches.

Location	Volume	Interpreted trigger	Minimum runout distance (km)	Maximum recorded transit velocity (m/s)	Velocity calculation based on
Submarine particulate density currents, including turbidity currents and volcaniclastic density currents					
1979 Nice Airport, Mediterranean (*58*)	0.008 km^3	Construction activity; slope failure during airport extension	120	7	Seafloor cable breaks
Gioia Canyon, Mediterranean (*59*)	Not known	Construction activity; slope failure near port	15	4.5	Seafloor cable breaks
1929 Grand Banks landslide, Newfoundland (*40*)	>200 km^3	Magnitude 7.2 earthquake triggered continental slope failure	800	19.1	Seafloor cable breaks
2006 Gaoping Canyon, Taiwan (*37*)	Not known	Magnitude 7.0 Pingtung earthquake	380	20	Seafloor cable breaks
1954 Orleansville earthquake, Algeria (*60*)	Not known	Magnitude 6.7 Orleansville earthquake	Not known	20.5	Seafloor cable breaks
2009 Gaoping Canyon, Taiwan (*37*)	Not known	Large river flood after typhoon	380	10.3	Seafloor cable breaks
2009 Gaoping Canyon, Taiwan (*37*)	Not known	Large river flood during typhoon	650	16.6	Seafloor cable breaks
2020 Deep-sea Congo Canyon, West Africa (*19*)	2.675 km^3	Low spring tide after large river flood	1130	8	Cable breaks and moored acoustic Doppler current profiler array
2015–17 Monterey Canyon, California (*55*)	Not known	Oceanographic trigger related to along-shelf transport	50	7.2	Moored acoustic Doppler current profiler array
2019–20 Whittard Canyon, northeast Atlantic (*61*)	Not known	Oceanographic trigger related to across-shelf transport	50	8	Moored acoustic Doppler current profiler array
2022 Hunga volcano, Tonga (this study)	>6.3 km^3 [based on deposited volume on slopes outside of the caldera (*15*)]	Eruption column collapse	>100	33.8	Cable breaks
Subaerial particulate density currents					
Largest snow avalanches (*22*, *62*)	0.01 km^3	Various	3 to 5	70	Radar and pressure plate measurements
Terrestrial pyroclastic density currents	Up to 5.5 km^3 for those with reported speeds, but can exceed hundreds of cubic kilometers	Dome or flank collapse or phreatomagmatic eruption	Up to tens of kilometers	7 to 210	See table S4 (*52*)

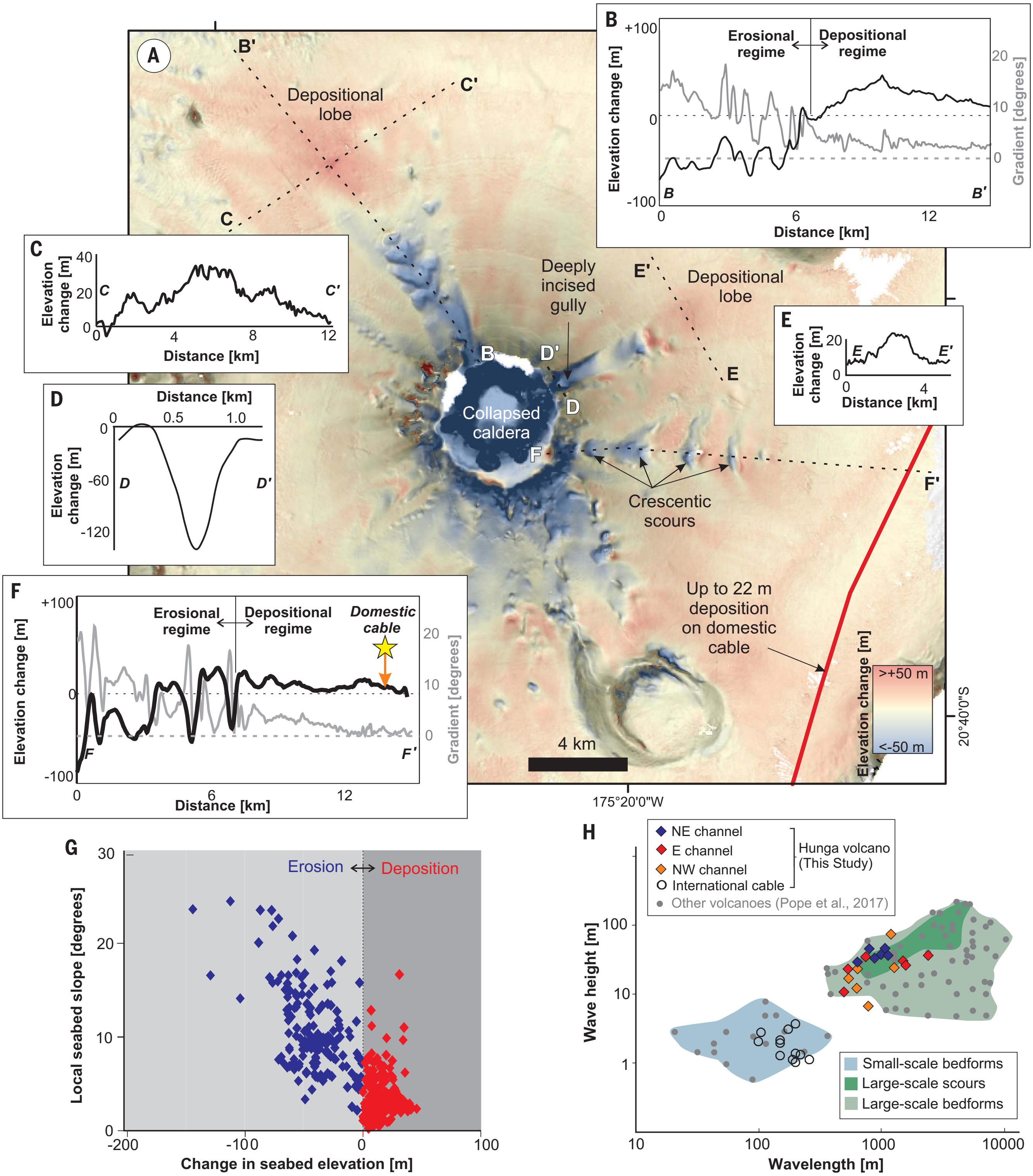


Fig. 2. Sculpting of the seafloor by powerful volcaniclastic density currents on the slopes proximal to Hunga volcano. (**A**) Elevation-difference map generated between pre- (2017) and post-eruption (April to June 2022) bathymetric surveys shows localized but major seafloor changes. The domestic cable is shown as a red line. (**B** to **F**) Elevation changes shown for selected locations in cross section, including the incision of deep (locally >120 m) gullies and upslope-migrating crescentic scours on steep (>10°) slopes [(B), (E), and (F)] and the deposition of thick (up to 40 m) lobes [(C) and (D)] where slopes shallow. (**G**) Cross plot of change in seabed elevation and local seafloor gradient (based on pre-eruption bathymetry) illustrating how erosion dominates on the steepest slopes, whereas deposition is largely restricted to slopes <10°. (**H**) Comparison of bedform morphometrics observed on the proximal flanks of Hunga volcano and around the area of the damage to the international cable with those from a database based on 17 volcanic islands worldwide (*8*) to show that the large-scale scours and bedforms plot within the range of those previously interpreted to relate to large explosive eruptions.

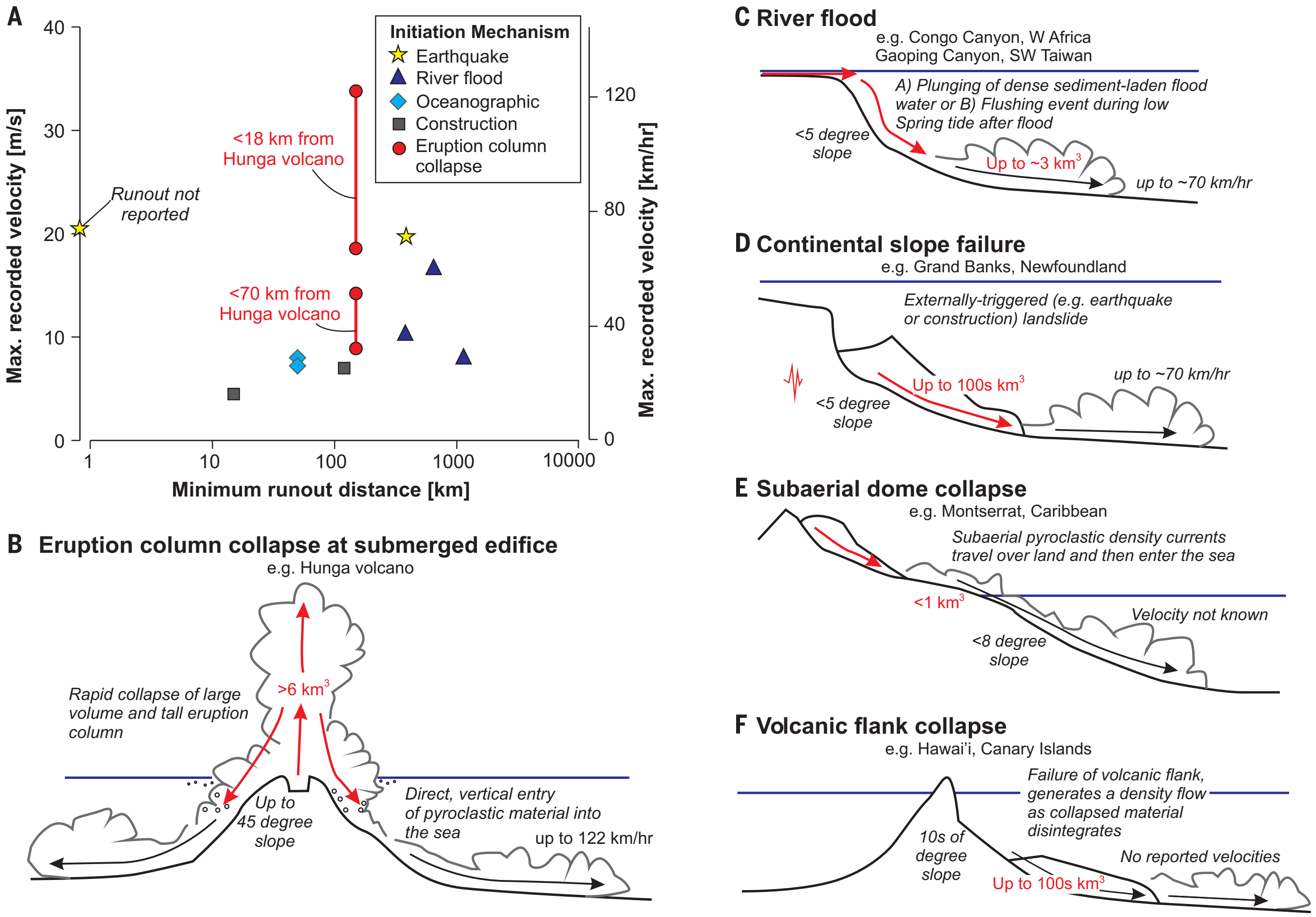


Fig. 3. Density currents triggered by the Hunga volcano eruption are the fastest reported for any submerged particulate density current to date. (**A**) Measured velocities and minimum runout distances for different underwater particulate density currents, categorized according to their triggers (as detailed in Table 1). Precise runout distances cannot be presented, as the current often ran out beyond the monitoring array or the location of seafloor cables. (**B** to **F**) Schematics illustrating the inception mechanisms for different submerged density currents: (B) rapid-eruption column collapse, causing vertical plunging of dense pyroclastic material onto an exceptionally steep slope, as seen at Hunga volcano; (C) river flood–triggered turbidity currents that enter the ocean either laterally (where sediment is flushed offshore) or obliquely (where dense sediment–laden flood water plunges), as observed in the Gaoping and Congo canyons; (D) continental slope collapses that are initiated by external ground disturbances, such as large earthquakes or construction activity, which can initiate on very low angle slopes; (E) initiated by subaerial volcanic dome collapse entering the ocean obliquely, as in the case of Montserrat; (F) initiated by the collapse of volcanic island flanks.

volcanic rock and/or coarse granular material. Deposition occurs as slope angles reduce (<10°), where well-defined lobes (up to 40 m thick and 7 km wide) accumulated downstream of two of the erosional chutes. As this lobate deposition occurs on relatively steep slopes, a dense granular current with high basal friction was likely responsible for the proximal depositional lobes (*41*–*43*). This is in stark contrast to turbidity currents, which typically do not deposit on steep slopes or, where they do, leave only very thin deposits (*44*). Indeed, some turbidity currents only deposit on slopes of less than 0.05° (*45*).

The intense erosion on steep slopes (Fig. 2) caused currents to increase their sediment mass substantially, thereby enhancing their power and mobility (*46*). For context, the eroded volume is 0.5 km^3 greater than the largest known historical volcanic landslide [3 km^3, Mount St. Helens (*47*)] and ~1 km^3 greater than the sediment volume eroded by the longest monitored turbidity current that traveled >1000 km along the deep-sea Congo Canyon (*19*). Stepped trains of upslope-migrating crescentic scours and large-scale bedforms on steep slopes evidence Froude-supercritical currents undergoing a series of hydraulic jumps (*48*–*50*). Similar scours and bedforms are a common feature proximal to the initiation of other large-volume particulate density currents, including those in nonvolcanic settings [e.g., the 1 km^3 earthquake-triggered turbidity current in Kaikōura Canyon (*51*)] and are of similar scale to those thought to diagnose past large explosive eruptions on submerged volcanic flanks (*25*–*28*) (Fig. 2H). We confirm this previously hypothesized link, showing that multiple radial chutes and bedforms can be formed during an individual explosive eruption, which has important implications for assessing hazards from seabed geomorphology at other submerged volcanoes worldwide.

Depositional distal regime

The remaining surveyed area is characterized by widespread, relatively featureless blanketed deposition, with an average of +2.8 m elevation change (i.e., net depositional). In the valley southeast of Hunga volcano (location of the domestic cable), slopes rapidly reduce to <1.5°, and ponded deposition occurred (up to 27 m thick). Even where elevation change was not resolvable from repeat seafloor surveys in deep water, sampling (up to 108 km from

the caldera) recovered volcaniclastic deposits emplaced by density currents (Fig. 1). Density current deposits are generally of decimeter thickness; comprise sand to granule–sized volcanic material that fines upward, with internal lamination, ripples, and sharp bases; and are geochemically linked to the 2022 Hunga volcanic eruption (*16*). Those deposits are overlain by thinner deposits interpreted as ash fallout (*52*) (fig. S2 and table S2). The area around the international cable was not covered by preexisting bathymetric data; however, detailed seafloor backscatter data reveal trains of bedforms (1 to 2 m wave height, 100 to 200 m wavelength) within a valley between seamounts (fig. S1). These bedforms evidence a complex flow pathway, as corroborated by modeling (*15*). So intense was the topographic steering, that the cable was moved 5 km toward the volcano by the density current (fig. S1).

Contrasting behavior of underwater volcaniclastic density currents

These depositional and erosional patterns contrast with observations from repeat seafloor surveys and sampling of terrestrially initiated pyroclastic density currents that entered the ocean offshore Montserrat in the Caribbean (*12–14*). Deposits offshore Montserrat formed in two parts: (i) coarse-grained ridges up to 60 m thick within 3 to 4 km of the ocean-entry point formed by a dense granular current; and (ii) a broader deep-water lobe, comprising centimeters-thick fine-grained deposits related to a dilute turbidity current (*12–14*). The proximal deposition of ridges offshore Montserrat contrasts with the erosional chutes, scours, and asymmetric bedforms we observe on Hunga volcano. Further, the lobes at Hunga volcano are much higher relief, thicker (tens of meters thick rather than centimeters thick), on steeper slopes (up to between 8° and 10° compared with <2°), and likely composed of coarser material, with deposits sampled at least 108 km from the edifice, compared with 40 km offshore Montserrat (Fig. 1). Seafloor cores show that coarse granular deposits were deposited at least 80 km from the Hunga edifice, indicating that currents maintained high densities over this distance, also in contrast with the dilute flow origin for distal finer-grained deposits offshore Montserrat (*12–14*).

Fast and long-runout underwater volcaniclastic density currents

Large volumes of pyroclastic material were delivered into the ocean during the eruption at Hunga volcano, creating dense underflows that were steered down its steep flanks. It is most likely that the pyroclastic material was predominantly derived from partial collapses of the eruption column, but we cannot fully preclude that currents may have been fed in part by other eruption processes, including jetting or fountaining. Initially, these volcaniclastic density currents would have been a dominantly gas-particle mixture, transitioning into a water-particle mixture as they cooled and mixed with seawater (*23*). We cannot discern the precise extent of that transition, underlining our broader use of "volcaniclastic" rather than "pyroclastic" density current. Volcaniclastic density currents were initially steered along preexisting relief into a valley 15 km southeast of the caldera. In this location, currents intersected with the domestic cable side-on and were deflected to the north and south (i.e., parallel to the cable) by the topography. On the basis of the time between the first collapses of the eruptive column into the ocean (T_{0a} and T_{0b}) and severing of the domestic cable, we calculate a distance-averaged transit (front) flow velocity of 63 to 122 km/hour (17.6 to 33.8 m/s; table S1). This observation is notable, given the inherent challenges in underwater monitoring, particularly during an ongoing large eruption.

Despite the higher resistance provided by the surrounding seawater compared with air, the velocities of volcaniclastic density currents offshore Hunga volcano fall within the range measured for pyroclastic density currents on land (*20*) (table S4). The velocities at Hunga volcano are higher than previously documented for underwater density currents elsewhere in the ocean, including turbidity currents triggered by large-magnitude earthquakes, river floods, and oceanographic processes (Table 1 and Fig. 3). The fastest transit velocities for turbidity currents are up to 72 km/hour [20 m/s; based on cable breaks during the 1929 Grand Banks landslide (*40*) and the 2006 Pingtung earthquake offshore Taiwan (*37*)]. The volcaniclastic density currents at Hunga volcano were steered into deeper water along tortuous paths created by irregular volcanic topography, where they severed the international cable 70 km from the volcano (Fig. 1). Assuming this current was the same one that also broke the domestic cable indicates a transit velocity of 32 to 51 km/hour (8.9 to 14.2 m/s). This is extraordinarily fast given the distance traveled. However, our data do not enable us to determine how many currents were triggered at Hunga volcano. It is plausible that damage to the international cable resulted from currents triggered by eruption collapses considerably later in the eruption cycle (after T_{0a} and conceivably after T_{0b}), in which case these distal velocities are underestimates.

What explains the fast current speeds?

The sheer mass and manner of delivery of material to the ocean (i.e., direct and rapid vertical entry of a fast-collapsing plume) at Hunga volcano is distinct from other mechanisms of particulate density current generation, such as river plumes that enter the ocean laterally, and landslide-triggered turbidity currents that initiate on far lower angle slopes and where material in the parent flow must first disintegrate and mix (Fig. 3). The dominantly downward rather than lateral trajectory toward and through the air-ocean boundary also makes the ocean-entry mechanism of dense pyroclastic material at Hunga volcano distinct from terrestrially initiated, ocean-entering pyroclastic density currents such as those observed at Montserrat. The pyroclastic density currents that entered the ocean at Montserrat first traveled 4 km laterally over land, along the Tar River valley, after a dome collapse (*12–14*). In contrast, the formative mechanism at Hunga volcano is better described as a vertical jet or fountain collapse of a gas-particle mixture (*53*), wherein a huge sediment load of dense volcanic pyroclasts [up to 2.8 g/cm^3 (*52*)] fell vertically from considerable height (several kilometers) as the eruption column collapsed into the ocean (Fig. 3).

According to the modified Chézy equation [a simplified approach often used to model behavior of turbidity currents (*52*)], to maintain the high current velocities observed at Hunga volcano would require a combination of a steep slope, thick current, and/or high sediment concentrations (expressed as the "depth averaged" value for a vertical profile through the current) (*52*, *54*). Assuming previously accepted basal and upper friction values for underwater density currents (*54*), to attain the high transit velocities on the edifice flanks requires current thicknesses on the order of tens of meters, with depth-averaged sediment concentrations of up to 5%, or else even-thicker currents (hundreds of meters thick) with lower depth-averaged sediment concentrations (1 to 2% concentration by volume). Near-bed sediment concentrations are likely substantially higher than these depth-averaged values (*8*, *55*). These concentrations are particularly high for underwater particulate density currents (*55*), as evidenced by the deposition of lobes on steep (up to ~10°) slopes of the edifice, implying a dense granular basal layer (*41–43*). Currents at Hunga volcano had sufficient inertia to flow upslope in some areas (and overtop relief of up to 860 m), such as south of the domestic cable break and to reach the international cable (figs. S4 and S5). This inertia was provided by their initial velocity and concentration and by additional entrained mass due to the substantial seafloor erosion. We conclude that fast velocities proximal to Hunga volcano result from (i) the potential energy generated from the direct, vertical entry of dense and large-volume pyroclastic fluxes into the ocean; (ii) the extremely steep (up to 45°) submerged edifice flanks, which were the location of the impact zone for the collapsed material (*53*); and (iii) the additional mass gained by the currents as they eroded into the edifice flanks.

Implications for hazards assessment

A push to enhance telecommunications links across the South Pacific and Caribbean will necessitate the crossing of volcanically active regions by new subsea cable routes with a need to assess threats posed to remote coastal communities and the international communications infrastructure that serves them. This requires more-extensive seafloor mapping to identify submerged volcanoes that may experience similar eruptions, while offshore monitoring, such as using fiber-optic sensing along telecommunications cables (*56*, *57*), is required to provide early warning. We show that volcaniclastic density currents triggered when an eruption collapses into the ocean can maintain high densities over distances of >100 km and attain speeds up to 122 km/hour, providing a fundamentally new view of their behavior and associated hazards. We confirm that bedforms observed on many other shallow submerged volcanoes worldwide can be produced by powerful eruptions, demonstrating that the hazards experienced at Hunga volcano can occur elsewhere (*4*, *8*, *27*). Explosive eruptions from these often unsurveyed and unmonitored submerged volcanoes can produce high-energy submarine density currents and warrant far greater consideration as tsunamigenic sources and as primary threats to vulnerable coastal communities and critical subsea infrastructure.

REFERENCES AND NOTES

1. G. Lube, E. C. Breard, T. Esposti-Ongaro, J. Dufek, B. Brand, *Nat. Rev. Earth Environ.* **1**, 348–365 (2020).
2. E. C. Breard *et al.*, *Nat. Geosci.* **9**, 767–771 (2016).
3. E. C. Breard, G. Lube, *Earth Planet. Sci. Lett.* **458**, 22–36 (2017).
4. C. E. J. de Ronde, G. J. Massoth, E. T. Baker, J. E. Lupton, in *Volcanic, Geothermal, and Ore-Forming Fluids: Rulers and Witnesses of Processes within the Earth*, S. F. Simmons, I. Graham, Eds., Special Publication No. 10 (Society of Economic Geologists, 2003), pp. 91–109.
5. M. R. Auker, R. S. J. Sparks, L. Siebert, H. S. Crosweller, J. Ewert, *J. Appl. Volcanol.* **2**, 2 (2013).
6. L. G. Mastin, J. B. Witter, *J. Volcanol. Geotherm. Res.* **97**, 195–214 (2000).
7. B. Sovilla, M. Schaer, M. Kern, P. Bartelt, *J. Geophys. Res. Earth Surf.* **113**, F01010 (2008).
8. E. L. Pope *et al.*, *Sci. Adv.* **8**, eabj3220 (2022).
9. B. Sovilla, J. N. McElwaine, M. Y. Louge, *C. R. Phys.* **16**, 97–104 (2015).
10. R. S. J. Sparks, H. Sigurdsson, S. N. Carey, *J. Volcanol. Geotherm. Res.* **7**, 87–96 (1980).
11. S. Carey, H. Sigurdsson, C. Mandeville, S. Bronto, in *Volcanic Hazards and Disasters in Human Antiquity*, Floyd W. McCoy, C. Heiken, Eds., Special Paper 345 (Geological Society of America, 2000), pp. 1–14.
12. J. Trofimovs *et al.*, *Geology* **34**, 549–552 (2006).
13. J. Trofimovs, R. S. J. Sparks, P. J. Talling, *Sedimentology* **55**, 617–634 (2008).
14. J. Trofimovs *et al.*, *Bull. Volcanol.* **74**, 391–405 (2012).
15. S. Seabrook *et al.*, Research Square [Preprint] (2023); https://www.researchsquare.com/article/rs-2395332/v1.
16. C. Gomez, F. Lavigne, D. S. Hadmoko, N. Lespinasse, P. Wassmer, *Geomorphology* **110**, 118–127 (2009).
17. D. E. Jessop *et al.*, *J. Volcanol. Geotherm. Res.* **245–246**, 81–97 (2012).
18. M. A. Hofton, E. Malavassi, J. B. Blair, *Geophys. Res. Lett.* **33**, L21306 (2006).
19. P. J. Talling *et al.*, *Nat. Commun.* **13**, 4193 (2022).
20. L. Scharff, M. Hort, N. R. Varley, *Sci. Rep.* **9**, 7386 (2019).
21. G. A. Valentine, *J. Volcanol. Geotherm. Res.* **87**, 117–140 (1998).
22. I. Vilajosana, G. Khazaradze, E. Surinach, E. Lied, K. Kristensen, *Cold Reg. Sci. Technol.* **49**, 2–10 (2007).
23. P. Kokelaar, P. Raine, M. J. Branney, *Bull. Volcanol.* **70**, 23–54 (2007).
24. P. Kokelaar, S. Koniger, *J. Geol. Soc. London* **157**, 517–536 (2000).
25. J. T. Gilchrist, A. M. Jellinek, E. E. Hooft, S. Wanket, *Nat. Geosci.* **16**, 446–453 (2023).
26. E. L. Pope *et al.*, *Earth Planet. Sci. Lett.* **493**, 12–24 (2018).
27. D. Casalbore *et al.*, *Sedimentology* **68**, 1400–1438 (2021).
28. J. Karstens *et al.*, *Earth Planet. Sci. Lett.* **616**, 118215 (2023).
29. J. C. Borrero *et al.*, *Pure Appl. Geophys.* **180**, 1–22 (2023).
30. P. Lynett *et al.*, *Nature* **609**, 728–733 (2022).
31. K. Pakoksung, A. Suppasri, F. Imamura, *Sci. Rep.* **12**, 15187 (2022).
32. P. Poli, N. M. Shapiro, *Geophys. Res. Lett.* **49**, e2022GL098123 (2022).
33. S. R. Proud, A. T. Prata, S. Schmauß, *Science* **378**, 554–557 (2022).
34. S. J. Purkis *et al.*, *Sci. Adv.* **9**, eadf5493 (2023).
35. D. A. Yuen *et al.*, *Earthq. Res. Adv.* **2**, 100134 (2022).
36. The World Bank and the Global Facility for Disaster Reduction and Recovery, "The January 15, 2022 Hunga Tonga-Hunga Ha'apai eruption and tsunami, Tonga: Global rapid post disaster damage estimation (grade) report" (International Bank for Reconstruction and Development/The World Bank, 2022); https://thedocs.worldbank.org/en/doc/b69af83e486aa652d4232276ad698c7b-0070062022/original/GRADE-Report-Tonga-Volcanic-Eruption.pdf.
37. L. Carter, R. Gavey, P. J. Talling, J. T. Liu, *Oceanography* **27**, 58–67 (2014).
38. J. Syvitski *et al.*, *Nat. Rev. Earth Environ.* **3**, 179–196 (2022).
39. M. Cassidy, L. Mani, *Nature* **608**, 469–471 (2022).
40. B. C. Heezen, D. B. Ericson, M. Ewing, *Deep Sea Res.* **1**, 193–202 (1954).
41. R. M. Iverson, *Rev. Geophys.* **35**, 245–296 (1997).
42. R. M. Iverson, J. W. Vallance, *Geology* **29**, 115–118 (2001).
43. J. J. Major, *J. Geol.* **105**, 345–366 (1997).
44. C. J. Stevenson, C. A. L. Jackson, D. M. Hodgson, S. M. Hubbard, J. T. Eggenhuisen, *J. Sediment. Res.* **85**, 1058–1081 (2015).
45. P. J. Talling *et al.*, *Nature* **450**, 541–544 (2007).
46. A. Manganet *et al.*, *J. Geophys. Res.* **115**, F03040 (2010).
47. O. Korup *et al.*, *Earth Planet. Sci. Lett.* **261**, 578–589 (2007).
48. M. J. Cartigny, D. Ventra, G. Postma, J. H. van Den Berg, *Sedimentology* **61**, 712–748 (2014).
49. W. O. Symons, E. J. Sumner, P. J. Talling, M. J. Cartigny, M. A. Clare, *Mar. Geol.* **371**, 130–148 (2016).
50. O. E. Sequeiros, *J. Geophys. Res. Oceans* **117**, C04003 (2012).
51. J. J. Mountjoy *et al.*, *Sci. Adv.* **4**, eaar3748 (2018).
52. See supplementary materials.
53. G. A. Valentine, *Bull. Volcanol.* **82**, 20 (2020).
54. K. Konsoer, J. Zinger, G. Parker, *J. Geophys. Res. Earth Surf.* **118**, 216–228 (2013).
55. C. K. Paull *et al.*, *Nat. Commun.* **9**, 4114 (2018).
56. P. Jousset *et al.*, *Nat. Commun.* **13**, 1753 (2022).
57. T. Nishimura *et al.*, *Sci. Rep.* **11**, 6319 (2021).
58. D. J. Piper, B. Savoye, *Sedimentology* **40**, 557–582 (1993).
59. P. Colantoni *et al.*, *Giornale di Geologia* **54**, 190–213 (1992).
60. B. C. Heezen, M. Ewing, *AAPG Bull.* **39**, 2505–2514 (1955).
61. M. S. Heijnen *et al.*, *Nat. Commun.* **13**, 3448 (2022).
62. D. McClung, P. Shaerer, *The Avalanche Handbook* (Mountaineers Books, ed. 3, 2006).
63. K. Mackay, Digital Elevation Models of Hunga Volcano; pre- and post- 15 January 2022 eruption, version 1.0, Zenodo (2022); https://doi.org/10.5281/zenodo.7456324.

ACKNOWLEDGMENTS

We thank the Kingdom of Tonga for allowing us to undertake this research and the Nippon Foundation-GEBCO Seabed 2030 project and their alumni for their support. This work would not have been possible without the captain, crew, and scientists aboard RV *Tangaroa* (Voyage TAN2206) and the use of SEA-KIT International's USV *Maxlimer* for mapping the caldera. We extend our gratitude to everyone involved in these voyages and for assistance in processing the results. We thank B. Sugar of South Sea Charters, Nuku'alofa, for providing the photographs of the eruption plume that appear in fig. S6. The staff of the British Ocean Sediment Core Research Facility (BOSCORF), including C. McGuire, R. Garnett, M. Charidemou, and M. Edwards, are thanked for supporting MSCL-CIS data collection and for providing logistical support to receive and store cores in the UK. We thank R. Williams and two anonymous reviewers for their constructive comments and suggestions. **Funding:** This work was supported by Natural Environment Research Council grant NE/X00239X/1 (M.A.C. and I.A.Y.); Natural Environment Research Council grant NE/X003272/1 (J.E.H., M.A.C., and I.A.Y.); Natural Environment Research Council grant NE/X002454/1 (D.T.); International Cable Protection Committee (M.A.C. and I.A.Y.); and the Nippon Foundation grant: NIWA-Nippon Foundation Tonga Eruption Seabed Mapping Project (S.W., R.W., S.S., K.M., E.L., and M.W.). **Author contributions:** Conceptualization: M.A.C., I.A.Y., P.J.T., E.P., K.M., R.W., S.W., T.K., S.H., C.d.R., M.U., S.K., S.F., S.P., D.V., R.R., and V.K. Methodology: K.M., M.W., I.A.Y., M.A.C., J.E.H., S.F., S.P., S.W., S.S., P.J.T., and E.P. Investigation: S.S., S.W., K.M., M.W., M.A.C., I.A.Y., J.E.H., S.C., and M.R. Visualization: M.A.C., I.A.Y., S.W., J.E.H., and S.C. Funding acquisition: M.A.C., I.A.Y., J.E.H., M.W., K.M., and D.T. Project administration: M.A.C., M.W., and K.M. Writing – original draft: M.A.C. and I.A.Y. Writing – review & editing: S.W., R.W., S.S., K.M., J.E.H., E.L., P.J.T., E.P., S.C., M.R., T.K., D.T., S.H., C.d.R., M.U., S.K., S.F., S.P., D.V., R.R., V.K., and M.W. **Competing interests:** M.A.C. is the marine environmental adviser to the International Cable Protection Committee. The other authors do not have any competing interests. **Data and materials availability:** Core logs, photographs, and coordinates are provided in the methods section of the supplementary materials. The pre- and post-eruption bathymetric data can be accessed in Zenodo (*63*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. This research was funded in whole or in part by the Natural Environment Research Council (NE/X00239X/1, NE/X003272/1, NE/X002454/1), a cOAlition S organization. The author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adi3038
Materials and Methods
Figs. S1 to S6
Tables S1 to S5
References (*64*–*89*)

Submitted 17 April 2023; accepted 1 August 2023
10.1126/science.adi3038

IMMUNOMETABOLISM

A type 2 immune circuit in the stomach controls mammalian adaptation to dietary chitin

Do-Hyun Kim[1], Yilin Wang[1], Haerin Jung[1], Rachael L. Field[1], Xinya Zhang[1], Ta-Chiang Liu[1], Changqing Ma[1], James S. Fraser[2], Jonathan R. Brestoff[1], Steven J. Van Dyken[1]*

Dietary fiber improves metabolic health, but host-encoded mechanisms for digesting fibrous polysaccharides are unclear. In this work, we describe a mammalian adaptation to dietary chitin that is coordinated by gastric innate immune activation and acidic mammalian chitinase (AMCase). Chitin consumption causes gastric distension and cytokine production by stomach tuft cells and group 2 innate lymphoid cells (ILC2s) in mice, which drives the expansion of AMCase-expressing zymogenic chief cells that facilitate chitin digestion. Although chitin influences gut microbial composition, ILC2-mediated tissue adaptation and gastrointestinal responses are preserved in germ-free mice. In the absence of AMCase, sustained chitin intake leads to heightened basal type 2 immunity, reduced adiposity, and resistance to obesity. These data define an endogenous metabolic circuit that enables nutrient extraction from an insoluble dietary constituent by enhancing digestive function.

Dietary fiber intake is associated with a lower risk of metabolic disorders such as obesity (*1*, *2*) and type 2 immune activation has been implicated in metabolic homeostasis (*3*–*5*), but little is known about how degradation of specific fibers influences host immunity and metabolism. In mammals, digestion is initiated in the upper gastrointestinal (GI) tract and is facilitated by mechanical forces, neural feedback, and enzymatic activities that coordinate chemical and physical disruption of the food bolus before passage into the highly absorptive small intestine. Digestion is essential for nutrient extraction, which is evident in bariatric surgical approaches that counteract overnutrition by reducing or bypassing gastric digestion (*6*). Nutrient availability is also limited by dietary fiber enrichment because most bulky insoluble polysaccharides are resistant to digestion by mammalian enzymes and undergo only limited degradation by distal gut microbes (*7*). A notable exception is chitin (β-1,4-poly-*N*-acetylglucosamine). One of the most abundant natural polysaccharides on Earth, chitin is a component of arthropods and fungi and is an initiator of type 2 immune responses. We hypothesized that chitin is digested in a distinctive manner because widely conserved chitinases are encoded by both commensal microbes and mammals, particularly those that consume chitin (*8*–*13*).

Dietary chitin induces gastric distension and type 2 immune triggering

Chitin activates lung group 2 innate lymphoid cells (ILC2s) by means of interleukin-25 (IL-25), IL-33, and thymic stromal lymphopoietin (TSLP) (*14*). To test GI responses to dietary chitin, "YRS" mice that express reporter alleles for ILC2 signature genes arginase-1 (Yarg; $Arg1^{YFP}$, where YFP is yellow fluorescent protein), IL-5 (Red5; $Il5^{tdTomato}$), and IL-13 (Smart13; $Il13^{hCD4}$) on wild-type (WT) and IL-25, IL-33 receptor (IL-33R), and thymic stromal lymphopoietin receptor (TSLPR) triple-knockout (TKO) (*15*) backgrounds were fed with standard chow containing either cellulose (control) or chitin as fiber (Fig. 1A and table S1). We tested 5 to 20% chitin, which approximates the dietary composition of insectivorous mammals (*11*, *12*). Food intake was similar, and GI transit time was unaffected by diet. However, we observed marked gastric distention and greater stomach contents in chitin-fed versus control-fed mice (Fig. 1B and fig. S1, A and B), indicating that dietary fiber influences stomach retention and stretch. Gastric epithelium rapidly

[1]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA. [2]Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA, USA.
*Corresponding author. Email: svandyken@wustl.edu

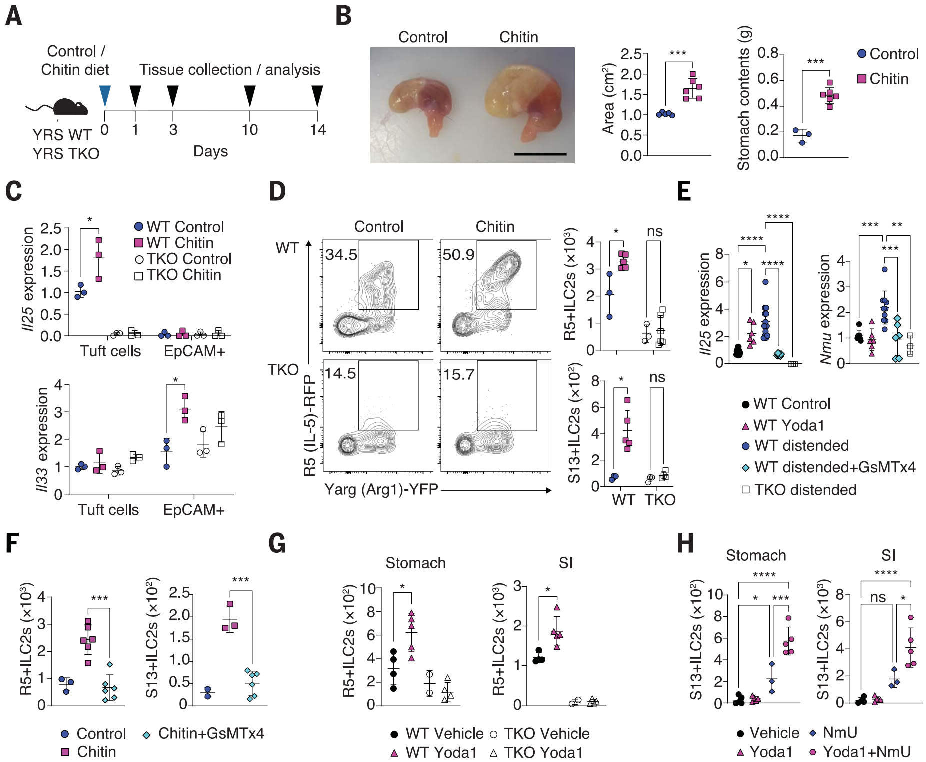


Fig. 1. Innate type 2 immune responses are triggered by gastric distension and dietary chitin. (**A**) Dietary responses in WT or TKO mice on triple-reporter (YRS) backgrounds. Representative stomach image (scale bar, 1 cm), stomach size, and luminal content (**B**); relative *Il25* and *Il33* expression in stomach tuft ($CD45^-EpCAM^+SiglecF^+$) and epithelial cells ($CD45^-EpCAM^+$) (**C**); and expression of R5 (*Il5*) and S13 (*Il13*) reporter alleles among stomach ILC2s ($Yarg^+$, pregated on $CD45^+Lin^-Thy1.2^+$) (**D**) in WT and TKO mice fed the indicated diet for 24 hours. RFP, red fluorescent protein. (**E**) Relative stomach gene expression in WT or TKO mice after Yoda1 administration or gastric distension. (**F**) R5 and S13 expression in stomach ILC2 after administration of the mechanosensitive ion-channel inhibitor GsMTx4 to mice fed the indicated diet for 12 hours. R5 (**G**) or S13 (**H**) expression in stomach and SI ILC2s after vehicle, Yoda1, and/or NmU administration. Data represent individual biological replicates except for the data in (C), which are pooled from two or three mice and are presented as means ± SD from two or more independent experiments ($n \geq 3$ mice per group). *P* values were calculated by unpaired *t* test [(B), (F), and (G)], one-way analysis of variance (ANOVA) [(E) and (H)], or two-way ANOVA with Tukey's multiple comparisons test [(C) and (D)]. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; ns, not significant.

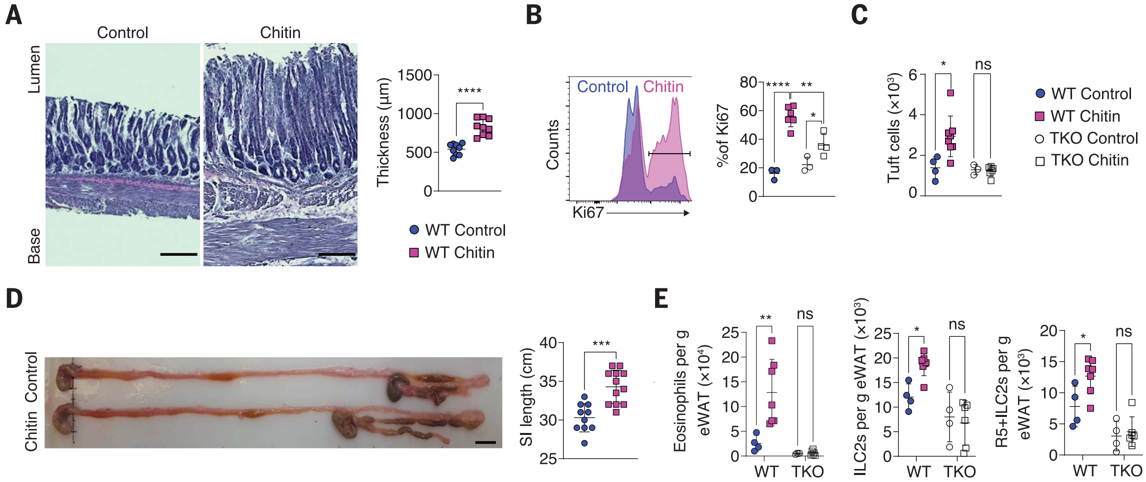


Fig. 2. Sustained chitin intake promotes GI remodeling and adipose ILC2 responses. Representative stomach histology images [hematoxylin and eosin (H&E) stained] (scale bars, 100 μm) (**A**), Ki67-expressing stomach epithelial cells (**B**), stomach tuft cells (**C**), and representative images of SI and small intestinal length (scale bar, 1 cm) (**D**) of the indicated mice fed the specified diet for 2 weeks. (**E**) Eosinophils, total ILC2s, and R5-expressing ILC2s per gram of epididymal white adipose tissue (eWAT) in WT and TKO mice fed the specified diet for 2 weeks. Data represent individual biological replicates and are presented as means ± SD from two or more independent experiments ($n \geq 3$ mice per group). P values were calculated by unpaired t test [(A), (C), and (D)], one-way ANOVA (B), or two-way ANOVA with Tukey's multiple comparisons test (E). *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; ns, not significant.

responded to chitin, inducing expression of the ILC2-activating cytokines *Il25* and *Il33* in stomach tuft cells and nontuft epithelial cells, respectively (Fig. 1C).

Dietary chitin increased stomach IL-5– and IL-13–producing ILC2s (Fig. 1D) and alternatively activated Arg1$^+$ macrophages, serum IL-5, and blood eosinophils, whose numbers expanded over time and in proportion with chitin content (fig. S1, C to E). Few IL-5– or IL-13–expressing stomach CD4$^+$ T cells were detected, however, and chitin responses were intact in *Rag1* knockout (*Rag1*-KO) mice, which lack B and T cells (fig. S1, F to I), consistent with predominantly innate immune activation. Chitin ingestion increased stomach expression of the gene encoding ILC2-activating neuropeptide neuromedin U (*Nmu*) (*16*, *17*) along with the gene encoding its receptor, *Nmur1*, among stomach-resident ILC2s (fig. S1, J and K). The expression of calcitonin gene–related peptide (CGRP), another ILC2-activating neuropeptide (*18*), was unaltered by dietary chitin (fig. S1L). Thus, gastric ILC2s may be synergistically activated by IL-25 and NMU, similar to intestinal ILC2s (*16*, *17*).

Dietary chitin also increased gastrin and glucagon-like peptide-1 (GLP-1), GI hormones that are induced by mechanical stretch after food ingestion (*19*). By contrast, serotonin, which responds to IL-33 (*20*), was unaffected by chitin (fig. S1, M and N). In addition, although IL-33 can be released by mechanical perturbation (*21*) and drives ILC2 responses to *Helicobacter pylori* infection and chemical injury (*22*, *23*), chitin-induced ILC2 accumulation and eosinophilia were unaffected in *Il1rl1*-KO mice (fig. S2, A and B). By contrast, ILC2 activation and cytokine production was abolished in TKO mice (Fig. 1, C and D), and eosinophilia was abrogated in both tuft cell–deficient *Pou2f3*-KO and TKO mice (fig. S2, C and D), whereas chitin-induced stomach distension was maintained. Thus, tuft cell–derived IL-25 appears to be the primary signal in gastric ILC2 responses to dietary chitin–induced stretch.

We inflated the stomach with air to recapitulate chitin-induced distension (fig. S2E). Within 2 hours, we observed increased expression of *Il25*, *Nmu*, and *Edn1*, which is induced by triggering the mechanosensitive ion channel Piezo1 (*24*, *25*). Conversely, in vivo administration of GsMTx-4, a stretch-activated channel inhibitor, blocked this response (Fig. 1E and fig. S2F) and abrogated ILC2 cytokine induction after chitin ingestion (Fig. 1F). Administration of the Piezo1 agonist Yoda1 induced ILC2 IL-5 production in WT but not TKO mice (Fig. 1G), suggesting that Piezo1 signaling activates gastric ILC2s through tissue cytokines, including IL-25. Accordingly, Yoda1 did not enhance gastric ILC2 responses before tuft cell development (fig. S2, G and H). *Nmur1* expression was reduced in TKO ILC2s compared with WT ILC2s (fig. S2I), which is consistent with prior reports (*16*) and suggests that gastric ILC2s acquire basal IL-5 expression and receptivity to multiple stretch signals during development. Combined Yoda1 and NMU administration also increased ILC2 IL-13 and KLRG1 expression compared with Yoda1 alone (Fig. 1H and fig. S2J). Thus, dietary chitin and mechanical stretch initiates type 2 immune responses that sensitize ILC2s to additional synergistic neuroimmune activating signals.

Sustained chitin intake promotes GI remodeling and adipose ILC2 responses

Dietary chitin remodeled GI tissues, inducing gastric epithelial proliferation, epithelial and submucosal thickening, and increased tuft cell abundance (Fig. 2, A to C). Proliferation was reduced in TKO mice compared with WT mice (Fig. 2B), indicating a requirement for IL-25, IL-33, and TSLP signaling in remodeling. Chitin lengthened the small intestine (SI), which was enriched with tuft cells, activated ILC2s, and eosinophils in WT but not TKO mice (Fig. 2D and fig. S3, A and B). These SI effects closely resembled those induced by helminths, protozoa (*Tritrichomonas* spp.), and increased luminal succinate (*26*–*29*). However, control and chitin-fed mice were *Tritrichomonas*-free, and cecal succinate levels were unaffected by chitin (fig. S3C). Thus, dietary chitin appears to initiate a distinctive type 2 immune circuit within the GI tract.

Chitin intake also stimulated type 2 immune responses in metabolically active tissues. Although lung ILC2s and eosinophils were unaffected by diet (fig. S3, D and E), visceral adipose from chitin-fed mice contained elevated numbers of eosinophils and IL-5–producing ILC2s (Fig. 2E). Inhibiting tissue lymphocyte egress by using the immunomodulator FTY720 reduced circulating ILCs (fig. S3F) but did not alter dietary chitin–induced eosinophil and ILC2 responses (fig. S3, G and H), suggesting that interorgan ILC2 migration did not mediate adipose effects (*30*). By contrast, adipose ILC2 activation and eosinophilia were abrogated in TKO mice (Fig. 2E), indicating that tissue-derived cytokines coordinate local ILC2 responses to chitin. Moreover, mice that lacked the shared signaling receptor for IL-4 and IL-13, IL4Rα, failed to induce stomach ILC2s, tuft cells, SI lengthening, adipose ILC2s, and eosinophils in response to chitin (fig. S3, I to M). ILC depletion in *Rag1*-KO mice impaired chitin-induced gastric ILC2 and tuft cell expansion (fig. S3, N and O), whereas tuft cells expanded normally in both *Il4*-KO and *Il5*-KO mice (fig. S3P), which supports a role for IL-13–producing ILC2s in chitin responses. Finally, GI tissue resilience, which relies on ILC2s and IL-13 in the absence of adaptive immunity (*31*), was enhanced in

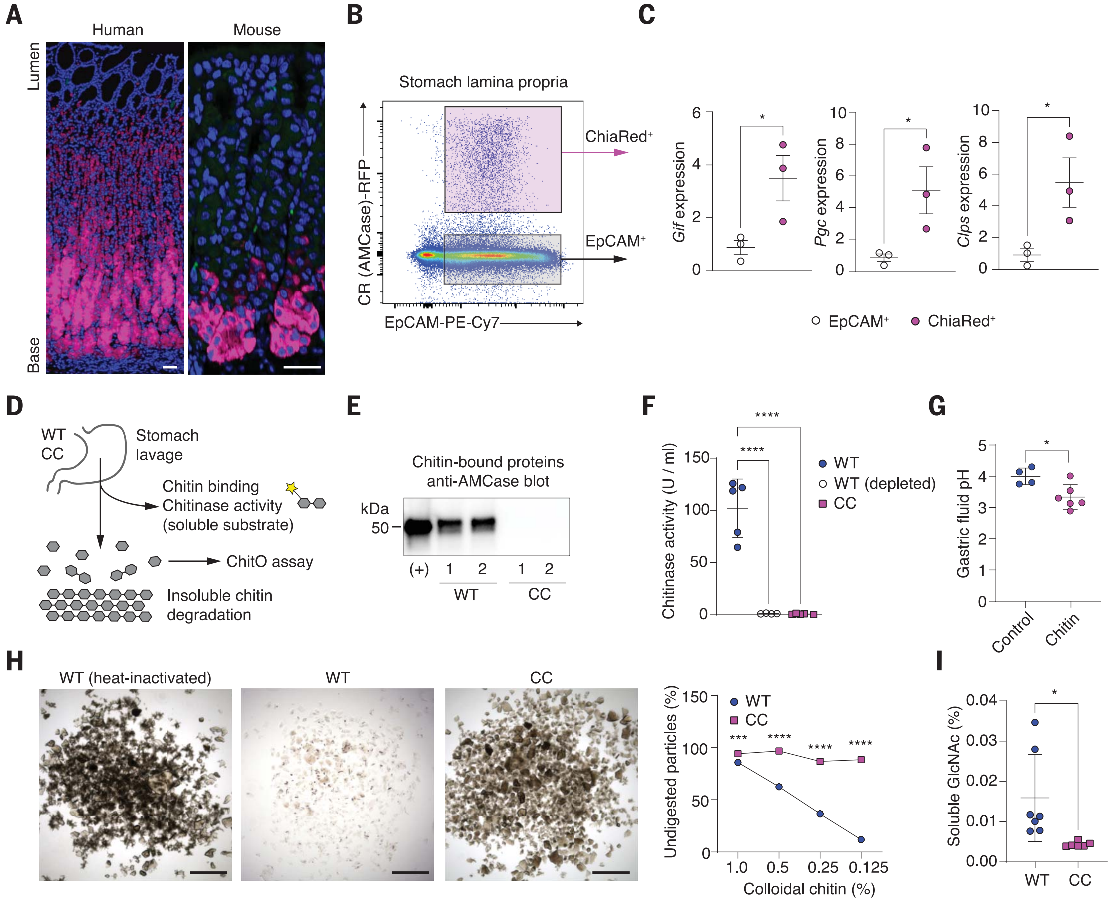


Fig. 3. AMCase is required for dietary chitin digestion. (**A**) Immunostaining of AMCase-expressing chief cells in glandular stomach. Magenta, AMCase; blue, 4′,6-diamidino-2-phenylindole (DAPI); green, autofluorescence (scale bars, 50 μm). Stomach ChiaRed$^+$ (CR: AMCase reporter; CD45$^-$EpCAM$^+$CR$^+$) or total (CD45$^-$EpCAM$^+$CR$^-$) epithelial cells were (**B**) isolated by fluorescence-activated cell sorting and (**C**) analyzed for relative expression of *Gif*, *Clps*, and *Pgc* by quantitative polymerase chain reaction (qPCR. (**D**) Analysis of chitin binding, digestion of soluble chitooligomer and insoluble crystalline chitin substrates, and production of GlcNAc reaction products by chito-oligomer oxidase (ChitO) assay in stomach lavage. (**E**) Immunoblot of chitin-bound AMCase proteins. The "(+)" indicates recombinant AMCase control. (**F**) Chitinase activity with soluble substrate in stomach lavage samples, with and without predepletion of AMCase by using insoluble chitin. (**G**) Gastric fluid pH in mice fasted overnight after 2 weeks on the indicated diet. (**H**) Digestion of insoluble colloidal chitin after 96-hour incubation with inactivated (heat-treated) or fresh stomach lavage from WT or CC mice (scale bars, 500 μm). (**I**) ChitO assay for soluble GlcNAc reaction products in supernatants from insoluble particle digestion in (H). Data points represent individual biological replicates except for the data in (C), which represent samples pooled from three or four mice. Data represent two or more independent experiments ($n \geq 3$ mice per group) and are presented as means ± SD. *P* values were calculated by unpaired *t* test. *$P < 0.05$; ***$P < 0.001$; ****$P < 0.0001$.

chitin-fed *Rag1*-KO mice infected with the helminth *Nippostrongylus brasiliensis* compared with control mice, as marked by increased ILC2s, serum IL-5, and eosinophils and improved worm expulsion (fig. S4, A to D).

Because dietary polysaccharides can be degraded by intestinal bacteria (*7*, *8*, *32*), we profiled the fecal microbiota from chitin- and control diet–fed mice. Mice maintained similar body weights regardless of diet (fig. S5A), which indicates equivalent nutrient extraction. However, chitin intake significantly enriched Bacteroidetes phyla, whereas Firmicutes were proportionally decreased in chitin-fed mice compared with control mice (fig. S5B and table S2). Thus, chitin alters GI bacterial composition, a finding that is consistent with prior work on dietary fiber enrichment (*8*).

We then tested whether type 2 immune responses to dietary chitin were dependent on commensal microbiota using germ-free (GF) and specific-pathogen–free (SPF) mice. GF and SPF mice maintained similar body weights regardless of diet, and dietary chitin induced gastric distension, SI lengthening, eosinophilia, ILC2 expansion, and tuft cell hyperplasia in both the GF and SPF conditions (fig. S5, C to G). These results indicated that dietary chitin induces innate type 2 immune responses independent of commensal microbes. To address possible developmental alterations in GF mice, we also depleted bacteria by administering antibiotics to adult SPF mice before dietary chitin intake. Consistent with GF results, stomach, SI, and adipose tissue chitin responses were unaffected by antibiotics (fig. S5, H to K). Thus, although chitin alters GI microbial composition and commensal microorganisms influence tuft cell succinate responses (*27*, *28*), dietary chitin initiates type 2 immune responses in the absence of commensal microbiota.

Acidic mammalian chitinase is required for dietary chitin digestion in mammals

Chitinases (EC 3.2.1.14) are widely conserved enzymes that cleave soluble chitooligomers and crystalline chitin substrates, liberating *N*-acetylglucosamine (GlcNAc). In contrast to insoluble dietary chitin fibers, GlcNAc did not induce gastric distension or GI type 2 immune responses (fig. S6, A to D), suggesting that undigested, insoluble chitin causes mechanical stretch and type 2 immune triggering. Chitinases can be expressed by microbes, including gut-resident bacteria (*8*). However, the lack of microbial involvement in chitin responses led us to consider acidic mammalian chitinase (AMCase), a mammalian chitinase secreted by respiratory epithelium, salivary glands, and stomach that has been evolutionarily linked to chitin consumption (*10*–*12*, *33*).

AMCase was expressed at the base of gastric glands in tissues from human sleeve-gastrectomy patients and 8-week-old mice (Fig. 3A), which is consistent with RNA sequencing (RNA-seq) data (fig. S7A) and prior reports (*10*, *33*–*35*), which supported chief cells as the main source of AMCase in the mammalian stomach. We further investigated AMCase-expressing cells using ChiaRed reporter mice, in which tdTomato and Cre recombinase are knocked into *Chia1* (which encodes AMCase). Homozygous ChiaRed (CC) mice lack AMCase (*36*). We lineage-traced AMCase-expressing cells by crossing ChiaRed with Rosa26-flox-stop-zsGreen [R26(LSL)-

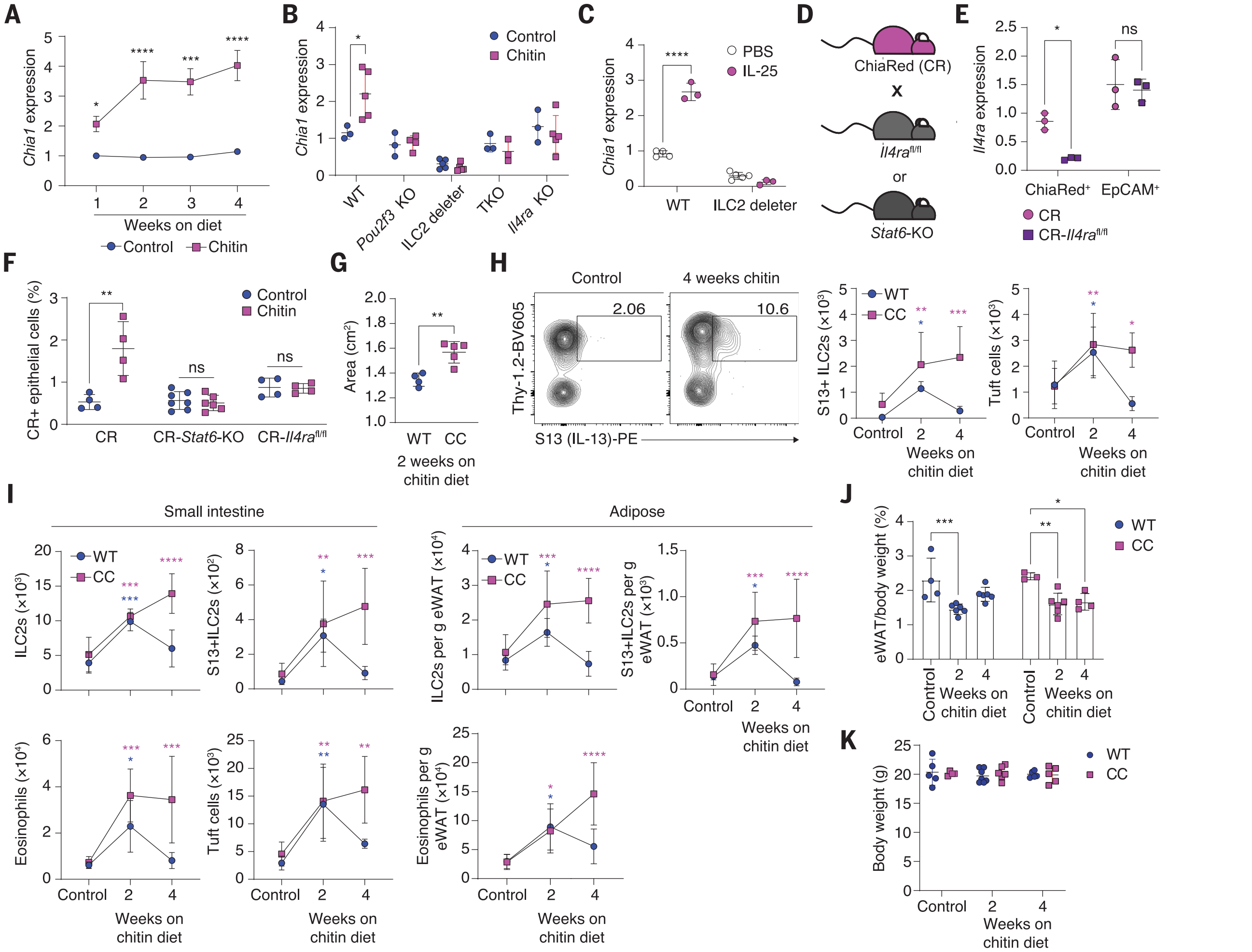


Fig. 4. Stomach adaptation to dietary chitin is controlled by a type 2 immune circuit. Relative *Chia1* stomach expression in WT (**A**) and indicated mouse strains (**B**) after 2 weeks on the indicated diet. (**C**) *Chia1* expression in stomach tissue from WT and ILC2-deficient deleter mice after IL-25 administration. (**D**) Breeding scheme to obtain CR-*Stat6*-KO, and CR-*Il4ra*$^{fl/fl}$ mice. (**E**) Relative *Il4ra* expression in ChiaRed$^+$ or total stomach epithelial cells (EpCAM$^+$) from ChiaRed and CR-*Il4ra*$^{fl/fl}$ mice. (**F**) Percentage of ChiaRed$^+$ (CR) chief cells out of total stomach epithelial cells in ChiaRed, CR-*Il4ra*$^{fl/fl}$, and CR-*Stat6*-KO mice fed the indicated diet for 2 weeks. (**G**) Stomach size of WT and CC mice fed the indicated diet for 2 weeks. S13 (IL-13)–expressing ILC2s and tuft cells in stomach (**H**); SI tuft cells, SI and adipose eosinophils, ILC2s, and S13$^+$ ILC2s (**I**); eWAT-to–body weight ratio (**J**), and body weights (**K**) of WT and CC mice fed control or chitin diets as indicated. Data represent individual biological replicates except for the data in (A), (E), (H), and (I), which are pooled from three to eight mice and are presented as means ± SD from two or more independent experiments ($n \geq 3$ mice per group). *P* values were calculated by unpaired *t* test [(A) to (C) and (E) to (G)] or two-way ANOVA with Tukey's multiple comparisons test [(H) to (K)]. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; ns, not significant.

zsGreen] mice. Consistent with antibody staining, AMCase-expressing ChiaRed$^+$zsGreen$^+$ cells were localized in gastric glands and concordantly coexpressed ChiaRed and zsGreen with age (fig. S7B), indicating that AMCase expression is sustained in terminally differentiated chief cells. ChiaRed$^+$ cells were also enriched for mature chief cell markers gastric intrinsic factor (*Gif*), colipase (*Clps*), and pepsinogen C (*Pgc*) (Fig. 3, B and C), which supports AMCase as a marker of zymogenic digestive cells.

We tested the contribution of AMCase to chitin digestion using GI luminal secretions from WT and CC (AMCase-deficient) mice. Chitinase activity was assessed on soluble chito-oligomers and crystalline chitin substrates (*37*), and chitin-binding proteins were isolated with magnetized chitin (Fig. 3D). WT stomach lavage contained AMCase that bound insoluble chitin (Fig. 3E) and exhibited robust chitinase activity that was absent in CC lavage and could be depleted by preincubation with chitin (Fig. 3F), which indicates that AMCase binds chitin and nonredundantly mediates stomach chitinase activity. Dietary chitin intake also reduced stomach pH (Fig. 3G), which enhances AMCase enzymatic activity (*10*, *33*) and suggests a coordinated physiological digestion response that involves acid-secreting parietal cells. Crystalline dietary chitin fibers were visibly digested and converted to soluble GlcNAc by WT stomach lavage, which reflects highly efficient chitinase digestive activity. However, this activity was absent in CC lavage, which retained undigested crystalline chitin particles and failed to produce soluble GlcNAc reaction products (Fig. 3, H and I). Neither control nor chitin chow elicited detectable epithelial ChiaRed or *Chia1* expression in lower GI tissues, including SI and cecum (fig. S8, A to C). However, WT SI lavage still contained chitinase activity that was absent in CC SI lavage, suggesting that AMCase produced in the stomach transits into the lower GI tract, where it also makes up the major source of chitinase activity (fig. S8D). Thus, although most insoluble dietary polysaccharides consumed by mammals resist digestion or undergo only limited degradation in the lower GI tract by commensal microbiota–derived glycosyl hydrolases (*7*, *8*), dietary chitin is primarily digested by host-encoded AMCase.

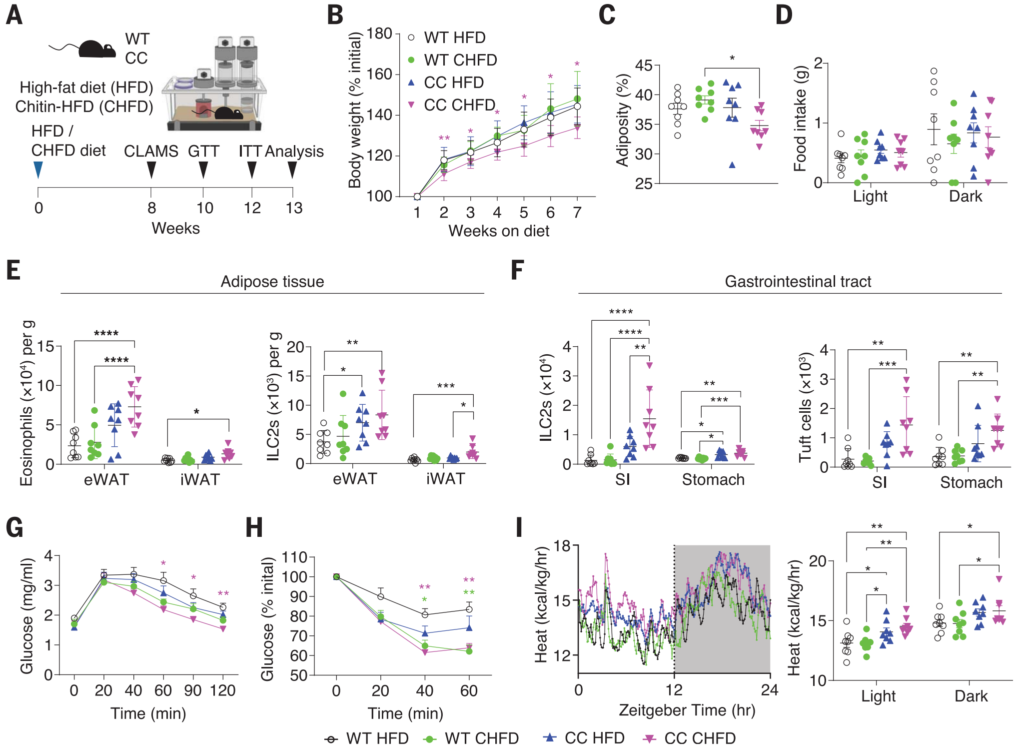


Fig. 5. Dietary chitin improves metabolic health in high-fat diet–induced obesity. (**A**) Experimental design. HFD- or CHFD-fed WT or CC mice were subjected to metabolic cage analyses (CLAMS), glucose tolerance tests (GTTs), and insulin tolerance tests (ITTs). [Part of the illustration was created with Biorender.com] Body weights (**B**), adiposity (**C**), and food intake (24-hour average) (**D**) of WT and CC mice after 8 weeks on the indicated diet. Eosinophils and ILC2s in adipose tissue (**E**) and tuft cells and ILC2s in SI and stomach tissues (**F**) in WT and CC mice after 13 weeks on the indicated diet. iWAT, inguinal white adipose tissue. (**G**) GTT curves at 10 weeks. (**H**) ITT curves at 12 weeks. (**I**) Heat curve and averages over 24 hours in CLAMS cages at 8 weeks. Data represent individual biological replicates except for data in (B), (E), and (F); in these panels, each data point represents eight pooled mice and are presented as means ± SD [(E) and (F)] or means ± SEM [(B) to (D) and (G) to (I)], from two or more independent experiments (n = 8 mice per group). P values were calculated by one-way ANOVA (C), two-way ANOVA [(E) and (F)], or two-way ANOVA with repeated measures [(B), (D), and (G) to (I)] with Tukey's multiple comparisons test. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$.

Stomach adaptation to dietary chitin is controlled by a type 2 immune circuit

Because lung AMCase expression is promoted by type 2 cytokines (*36*, *38*), we tested whether stomach AMCase is similarly regulated. Indeed, sustained dietary chitin intake increased *Chia1* expression in WT but not *Il4ra*-KO, *Pou2f3*-KO, ILC2-deleter mice that lack ILC2s (*15*), or TKO mice (Fig. 4, A and B). Tuft cell–derived IL-25, ILC2s, and IL-4Rα signaling were therefore implicated as major drivers of gastric AMCase expression. Type 2 triggering was recapitulated by IL-25 administration, which stimulated IL-13 production by ILC2s and stomach *Chia1* expression in WT but not ILC2-deficient mice (Fig. 4C and fig. S8E). Thus, ILC2-derived IL-13 is implicated in the expansion of AMCase-expressing chief cells.

To test this further, we crossed ChiaRed with *Stat6*-KO (CR-*Stat6*-KO) and *Il4ra*$^{fl/fl}$ mice (*39*) (CR-*Il4ra*$^{fl/fl}$), which enabled specific deletion of IL4Rα from AMCase-expressing cells (Fig. 4D). *Il4ra* was reduced in ChiaRed$^+$ stomach epithelial cells from CR-*Il4ra*$^{fl/fl}$ mice compared with ChiaRed controls, indicating successful Cre-mediated excision (Fig. 4E). Dietary chitin increased AMCase-expressing chief cells in WT ChiaRed mice but failed to occur in both CR-*Stat6*-KO and CR-*Il4ra*$^{fl/fl}$ mice (Fig. 4F), indicating that dietary chitin promotes chief cell AMCase expression through cell-intrinsic IL4Rα signaling. We also examined acute gastric injury and transient chief cell depletion by high-dose tamoxifen treatment (HDT), which models aspects of human atrophic corpus gastritis and AMCase loss due to *H. pylori* infection (*40*, *41*). After HDT, WT mice activated gastric ILC2s coincident with chief cell recovery, whereas TKO mice failed to recover *Chia1* (fig. S9, A and B). Thus, restoration of gastric homeostasis and chitin digestive capacity after epithelial injury depends on type 2 circuit activation.

These data suggest that mammals adapt to dietary chitin by inducing endogenous stomach type 2 immune responses to boost AMCase production. Accordingly, CC mice failed to reduce gastric distension compared with WT mice after 2 weeks of chitin intake (Fig. 4G). Stomach ILC2 activation and tuft cell hyperplasia were also sustained in CC versus WT mice over several weeks (Fig. 4H), which reflects unresolved circuit activation without AMCase-catalyzed chitin digestion. Consistent with improved digestion, WT mice attenuated distal type 2 immune triggering in the SI and adipose tissues. CC mice, by contrast, exhibited enhanced and prolonged type 2 triggering, characterized by greater SI tuft cell abundance, increased eosinophils, and increased IL-13–producing ILC2s in SI and adipose tissues after 4 weeks of chitin intake (Fig. 4I). Adipose tissue weight was also reduced in proportion to body weight in CC mice compared with WT mice, whereas body weights were similar (Fig. 4, J and K). Thus, both GI and adipose tissue homeostasis are regulated by the AMCase-mediated adaptation to dietary chitin.

Dietary chitin improves metabolic health in obesity

We next tested the impact of dietary chitin on obesity, which is influenced by neuronal-ILC2 interactions and type 2 cytokines (*3*–*5*, *42*). We fed WT and CC mice isocaloric high-fat diets containing either cellulose (control; HFD) or chitin (CHFD) fiber (Fig. 5A). Food intake was similar among all groups, and HFD-fed mice showed comparable body weight gain. However, CHFD-fed CC mice gained significantly less weight, with reduced adiposity and fat mass compared with WT mice (Fig. 5, B to D, and fig. S10A). Resistance to obesity in CHFD-fed CC mice was accompanied by adipose ILC2 and eosinophil accumulation (Fig. 5E), which have been previously linked with metabolic homeostasis (*3*–*5*), suggesting that altered dietary chitin digestion could modulate metabolism. Indeed, numbers of ILC2s and tuft cells were elevated in the stomach and SI tissues of CHFD-fed CC compared with those of WT mice (Fig. 5F), which reflects sustained type 2 triggering and suggests that AMCase-mediated dietary chitin digestion contributes to metabolic homeostasis.

Both dietary chitin and AMCase influenced metabolism in the context of high-fat diets because CHFD-fed WT and CC mice exhibited significantly improved insulin sensitivity compared with HFD-fed WT mice (Fig. 5, G and H). CHFD-fed CC mice exhibited lower fasting glucose compared with WT mice (fig.

S10B), which is consistent with differences in body weight and suggests that AMCase activity influences glucose homeostasis after chitin ingestion. Additionally, light-phase energy expenditure was increased in CC mice, and the respiratory exchange ratio was increased after CHFD feeding compared with HFD feeding in both WT and CC mice, despite no differences in core body temperature or activity (Fig. 5I and fig. S10, C to E), which is consistent with sustained type 2 immune triggering. Indeed, tuft cell–deficient *Pou2f3*-KO mice failed to exhibit CHFD-induced effects on insulin sensitivity (fig. S11, A to G), suggesting that type 2 circuit initiation is required for some metabolic aspects of chitin in a high-fat diet. Thus, disruption of the mammalian stomach's adaptation to dietary chitin alters nutrient uptake and metabolic homeostasis, which manifests in obesity.

Discussion

Chitin consumption has been linked to *CHIA* gene selection throughout primate evolution (*11*, *12*), and chitin-rich fungi and arthropods are constituents of the diets of both modern and ancient human populations (*13*). Mammalian glycosyl hydrolase gene selection has also been linked with starch consumption (*43*, *44*), but adaptations to specific dietary fibers after ingestion are mainly ascribed to shifts in gut microbial composition. As shown here, mammals encode an endogenous circuit that enables chitin catabolism and nutrient extraction through AMCase production. This pathway is triggered by gastric distension, neuropeptide release, and type 2 cytokine production caused by insoluble chitin fibers, which result in GI remodeling and chief cell AMCase induction over time. This in turn enables enhanced chitin digestion. In humans, *CHIA* variants resulting in lower chitinase activity are linked with asthma (*45*, *46*), which supports dual roles for AMCase in mucosal defense and nutrient extraction, as proposed initially (*33*). AMCase-producing chief cells produce additional digestive enzymes such as pepsinogen and lipase, which suggests that gastric ILC2s may coordinate a response that improves overall digestion of recalcitrant insect or fungal foods, perhaps representing a strategy for omnivores to adapt to varied diets. Although microbial composition is altered in response to chitin and other polysaccharides, we show that the mammalian adaptation does not rely on commensal microbiota and that AMCase is the primary source of chitinase activity in the GI tract, thereby distinguishing chitin digestion from other abundant insoluble polysaccharides such as cellulose. Our results further elucidate a mechanism for how chitin initiates type 2 immune initiation by means of mechanical stretch, thus connecting physical tissue perturbation with a branch of immunity increasingly recognized to maintain homeostasis in response to a wide variety of environmental disruptions as well as neural and dietary fluctuations. Intriguingly, the mammalian adaptation to chitin influences innate resistance to helminth infection and metabolic homeostasis, suggesting that chitin digestive pathways coevolved with intestinal helminths and may represent a therapeutic target in metabolic diseases such as obesity.

REFERENCES AND NOTES

1. J. L. Slavin, *Nutrition* **21**, 411–418 (2005).
2. S. K. Gill, M. Rossi, B. Bajka, K. Whelan, *Nat. Rev. Gastroenterol. Hepatol.* **18**, 101–116 (2021).
3. D. Wu *et al.*, *Science* **332**, 243–247 (2011).
4. A. B. Molofsky *et al.*, *J. Exp. Med.* **210**, 535–549 (2013).
5. J. R. Brestoff *et al.*, *Nature* **519**, 242–246 (2015).
6. N. T. Nguyen, J. E. Varela, *Nat. Rev. Gastroenterol. Hepatol.* **14**, 160–169 (2017).
7. H. J. Flint, K. P. Scott, S. H. Duncan, P. Louis, E. Forano, *Gut Microbes* **3**, 289–306 (2012).
8. J. L. Sonnenburg *et al.*, *Science* **307**, 1955–1959 (2005).
9. M. G. Paoletti, L. Norberto, R. Damini, S. Musumeci, *Ann. Nutr. Metab.* **51**, 244–251 (2007).
10. M. Ohno *et al.*, *Sci. Rep.* **6**, 37756 (2016).
11. C. A. Emerling, F. Delsuc, M. W. Nachman, *Sci. Adv.* **4**, eaar6478 (2018).
12. M. C. Janiak, M. E. Chaney, A. J. Tosi, *Mol. Biol. Evol.* **35**, 607–622 (2018).
13. M. C. Wibowo *et al.*, *Nature* **594**, 234–239 (2021).
14. S. J. Van Dyken *et al.*, *Immunity* **40**, 414–424 (2014).
15. S. J. Van Dyken *et al.*, *Nat. Immunol.* **17**, 1381–1387 (2016).
16. V. Cardoso *et al.*, *Nature* **549**, 277–281 (2017).
17. A. Wallrapp *et al.*, *Nature* **549**, 351–356 (2017).
18. P. Sui *et al.*, *Science* **360**, eaan8546 (2018).
19. T. Zhang, M. H. Perkins, H. Chang, W. Han, I. E. de Araujo, *Cell* **185**, 2478–2494.e28 (2022).
20. Z. Chen *et al.*, *Immunity* **54**, 151–163.e6 (2021).
21. R. Kakkar, H. Hei, S. Dobner, R. T. Lee, *J. Biol. Chem.* **287**, 6941–6948 (2012).
22. N. Satoh-Takayama *et al.*, *Immunity* **52**, 635–649.e4 (2020).
23. A. R. Meyer *et al.*, *Gastroenterology* **159**, 2077–2091.e8 (2020).
24. A. G. Solis *et al.*, *Nature* **573**, 69–74 (2019).
25. Y. Oh *et al.*, *Neuron* **109**, 1979–1995.e6 (2021).
26. J. von Moltke, M. Ji, H.-E. Liang, R. M. Locksley, *Nature* **529**, 221–225 (2016).
27. C. Schneider *et al.*, *Cell* **174**, 271–284.e14 (2018).
28. W. Lei *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 5552–5557 (2018).
29. M. S. Nadjsombati *et al.*, *Immunity* **49**, 33–41.e7 (2018).
30. Y. Huang *et al.*, *Science* **359**, 114–119 (2018).
31. A. E. Price *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11489–11494 (2010).
32. P. J. Turnbaugh *et al.*, *Nature* **444**, 1027–1031 (2006).
33. R. G. Boot *et al.*, *J. Biol. Chem.* **276**, 6770–6778 (2001).
34. M. Suzuki *et al.*, *J. Histochem. Cytochem.* **50**, 1081–1089 (2002).
35. X. Han *et al.*, *Cell* **172**, 1091–1107.e17 (2018).
36. S. J. Van Dyken *et al.*, *Cell* **169**, 497–509.e13 (2017).
37. B. A. Barad *et al.*, *Protein Sci.* **29**, 952–977 (2020).
38. T. A. Reese *et al.*, *Nature* **447**, 92–96 (2007).
39. D. R. Herbert *et al.*, *Immunity* **20**, 623–635 (2004).
40. W. J. Huh *et al.*, *Gastroenterology* **142**, 21–24.e7 (2012).
41. I. Nookaew *et al.*, *BMC Med. Genomics* **6**, 41 (2013).
42. F. Cardoso *et al.*, *Nature* **597**, 410–414 (2021).
43. G. H. Perry *et al.*, *Nat. Genet.* **39**, 1256–1260 (2007).
44. E. Axelsson *et al.*, *Nature* **495**, 360–364 (2013).
45. M. A. Seibold *et al.*, *J. Biol. Chem.* **284**, 19650–19658 (2009).
46. K. Okawa *et al.*, *Mol. Biol. Evol.* **33**, 3183–3193 (2016).
47. D. H. Kim *et al.*, A type 2 immune circuit in the stomach controls mammalian adaptation to dietary chitin, version 1. Dryad (2023); https://doi.org/10.5061/dryad.12jm63z3m.

ACKNOWLEDGMENTS

We thank K. Ravichandran for comments on the manuscript; J. Mills, R. Locksley, H.-E. Liang, J. Kipnis, M. Baldridge, C. Wilen, and C. Hsieh for advice and reagents; L. Mosby for expert technical assistance; and J. M. White at the Department of Pathology and Immunology Gnotobiotic Facility at Washington University in St. Louis for assistance with GF experiments. **Funding:** This work was supported in part by the Bursky Center for Human Immunology and Immunotherapy Programs, Center for Cellular Imaging, and Rheumatic Diseases Research Resource-based Center (NIH P30 AR073752) at Washington University in St. Louis; NIH DP5 OD028125 (J.R.B.); Burroughs Wellcome Fund CAMS #1019648 (J.R.B.); NIH R01HL148033, R01AI176660, and R21AI163640 (S.J.V.D.); and the National Research Foundation of Korea Award NRF-2020R1A6A3A03037855 (D.-H.K.). **Author contributions:** D.-H.K. and S.J.V.D. designed the study and performed experiments. Y.W. and H.J. performed experiments and provided advice. R.L.F., X.Z., and J.R.B. performed metabolic cage experiments, analyzed data, and provided expertise with metabolic studies. C.M. and T.-C.L. provided human stomach specimens and expertise, and J.S.F. provided biochemical reagents and expertise. All authors edited the final manuscript. S.J.V.D. directed the studies and wrote the manuscript with D.-H.K. **Competing interests:** J.R.B. is on the scientific advisory board of LUCA Science, Inc. The other authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or supplementary materials. 16*S* sequencing data are available in Dryad (*47*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add5649
Materials and Methods
Figs. S1 to S11
Tables S1 to S3
References (*48*–*52*)
MDAR Reproducibility Checklist

Submitted 20 June 2022; resubmitted 8 May 2023
Accepted 8 August 2023
10.1126/science.add5649

MEMBRANES

Carbon-doped metal oxide interfacial nanofilms for ultrafast and precise separation of molecules

Bratin Sengupta[1]†, Qiaobei Dong[1]†, Rajan Khadka[2], Dinesh Kumar Behera[1], Ruizhe Yang[3], Jun Liu[3], Ji Jiang[4], Pawel Keblinski[2], Georges Belfort[4], Miao Yu[1]*

Membranes with molecular-sized, high-density nanopores, which are stable under industrially relevant conditions, are needed to decrease energy consumption for separations. Interfacial polymerization has demonstrated its potential for large-scale production of organic membranes, such as polyamide desalination membranes. We report an analogous ultrafast interfacial process to generate inorganic, nanoporous carbon-doped metal oxide (CDTO) nanofilms for precise molecular separation. For a given pore size, these nanofilms have 2 to 10 times higher pore density (assuming the same tortuosity) than reported and commercial organic solvent nanofiltration membranes, yielding ultra-high solvent permeance, even if they are thicker. Owing to exceptional mechanical, chemical, and thermal stabilities, CDTO nanofilms with designable, rigid nanopores exhibited long-term stable and efficient organic separation under harsh conditions.

Industrially relevant molecular separations, for example, in pharmaceutical (*1*), petroleum (*2*), and chemical industries (*3*), involve harsh solvents at high temperatures. Ease of fabrication into thin films, especially through interfacial polymerization, makes polymers promising for scalable membrane fabrication (*4*). However, few polymeric membranes target these applications involving challenging industrially relevant conditions (*5*–*7*). Some recent breakthroughs have shown polymer stability at higher temperatures (*8*) but most suffer from instability under such severe conditions, undergoing aging and/or pore collapse due to polymer chain relaxation (*9*). Nonpolymeric counterparts such as carbon and graphene/graphene oxide (*3*, *10*–*14*), metal/covalent organic frameworks (MOF/COF) (*15*, *16*), and ceramics (*17*, *18*) having stable rigid pores are ideal for extending membrane separation to these harsh industrial conditions. While disorders/defects and intercrystalline grain boundaries are commonplace in graphene and MOF/COF membranes, causing reproducibility issues and challenging scale-up (*19*), ceramic membranes prepared by the traditional sol-gel method lack precise nanopore size control and have a thickness in the micrometer range. Materials that have precisely tunable rigid pores, while owning processibility of polymers, easily form defect-free continuous membranes and have excellent chemical, mechanical, and thermal stabilities of inorganic materials, are therefore missing. Interfacial reactions, analogous to that used for commercial polyamide desalination membrane fabrication (*20*), to generate amorphous defect-free inorganic nanofilms are highly desirable and might fill this material gap, extending membrane separation to harsher conditions.

[1]Department of Chemical and Biological Engineering and RENEW Institute, University at Buffalo, Buffalo, NY 14260, USA. [2]Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA. [3]Department of Mechanical and Aerospace Engineering, University at Buffalo, Buffalo, NY 14260, USA. [4]Howard P. Isermann Department of Chemical and Biological Engineering and the Center of Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180, USA.
***Corresponding author. Email: myu9@buffalo.edu**
†These authors contributed equally to this work.

Besides stability and selectivity, high permeance is critical, considering the volume of solvents processed in industry (*5*, *11*). Thickness reduction is an effective way of increasing permeance. Most recent studies seek selective layer thickness reduction, sometimes down to atomic thinness, such as ultrathin polymeric coatings (*5*, *21*, *22*) and monolayer graphene (*12*, *23*), to achieve high permeance. Although effective, this increases the probability of introducing defects and pinholes and thus likely results in scale-up challenges (*24*). We envision high-density nanopores with low tortuosity as another effective way of increasing permeance substantially. Although it is challenging to increase pore density without merging small nanopores into larger ones, this strategy can enhance membrane permeance, avoiding pitfalls of pushing membrane thickness to their thinnest limit.

Using a self-terminating interfacial reaction between metallic and organic reactants and subsequent calcination, we fabricated defect-free, continuous nanoporous films on different porous supports. Simple adjustment to synthesis and calcination conditions allowed precise tuning of these nanopores, at a molecular weight cut-off (MWCO) step change as small as ~100 g mol^{-1}, to prepare a series of organic solvent nanofiltration (OSN) or solvent-resistant nanofiltration membranes.

Porous nanofilms through interfacial reaction

Molecular layer deposition (MLD) is a vapor phase deposition technique for polymeric or organo-metallic hybrid films (*25*). Because of the layer-by-layer growth feature and nonselective deposition on all exposed surfaces, MLD is not appropriate for the ultrafast deposition of an ultrathin skin layer/separation membrane on the porous supports (*26*). Inspired by the self-limiting reactions of MLD, we used titanium tetrachloride ($TiCl_4$) and ethylene glycol (EG), as metallic and organic reactants, respectively, and developed an interfacial reaction process (Fig. 1A, I and II) to prepare a dense skin layer (Fig. 1A, III). Independent of support morphology—flat anodic aluminum oxide (AAO) or cylindrical α-alumina hollow fiber (HF)—we fill the support pores with liquid EG, followed by reaction with $TiCl_4$ at the pore mouth to form a nonporous organometallic hybrid film (OHF) (figs. S1 and S2). This generates an OHF skin layer at a rate 2 to 3 orders of magnitude faster (only 1 min) than the layer-by-layer MLD process. Upon carbon removal by thermal treatment/calcination, nanoporous carbon-doped metal oxide (CDTO) is generated (Fig. 1A, IV). We optimized fabrication conditions for rapid and defect-free synthesis of OHF (Supplementary Note 1, figs. S3 to S6, and tables S1 and S2). OHF can be formed on supports having different pore sizes by optimizing the $TiCl_4$ phase (vapor or liquid), concentration, and reaction time, as indicated by gas permeance and scanning electron microscopy (SEM) images. Reaction between liquid EG and $TiCl_4$ vapor at higher temperatures (150°C) forms the thinnest, defect-free OHF in the shortest time. Results for this OHF are reported henceforth. Undetectable permeance of N_2 (<10^{-12} mol m^{-2} s^{-1} Pa^{-1}) and methanol (<0.05 L m^{-2} h^{-1} bar^{-1}) measured using a commercial or a home-built permeation cell (figs. S7 and S8) indicate defect-free OHF formation within 1 min on AAO and 5 min on α-alumina HF. SEM images show evolution of OHF into a continuous, amorphous (confirmed by x-ray diffraction) nanofilm on a 50-nm porous HF (figs. S9 and S10, respectively).

We modeled this hybrid material using the large-scale atomic/molecular massively parallel simulator (LAMMPS) to understand pore generation during calcination (*27*) where the energy minimized structure was subjected to heat under an N_2 or O_2 atmosphere (fig. S11). Figure 1B shows the material formed after heat treatment in N_2 and O_2, respectively. Densely packed pores are generated throughout the material, and the resulting porous structure (fig. S2) is highly dependent on its final carbon content (Fig. 1B, I and II). Pore formation is comparatively faster in O_2 and accelerated by temperature (figs. S12 to S14). The surface area, pore volume, and porosity of CDTO are highly correlated with carbon retained after calcination, namely "carbon doping", in the titanium oxide network (Fig. 1B, III). Carbon partially leaves the hybrid material, and carbon removal is responsible for the pore formation

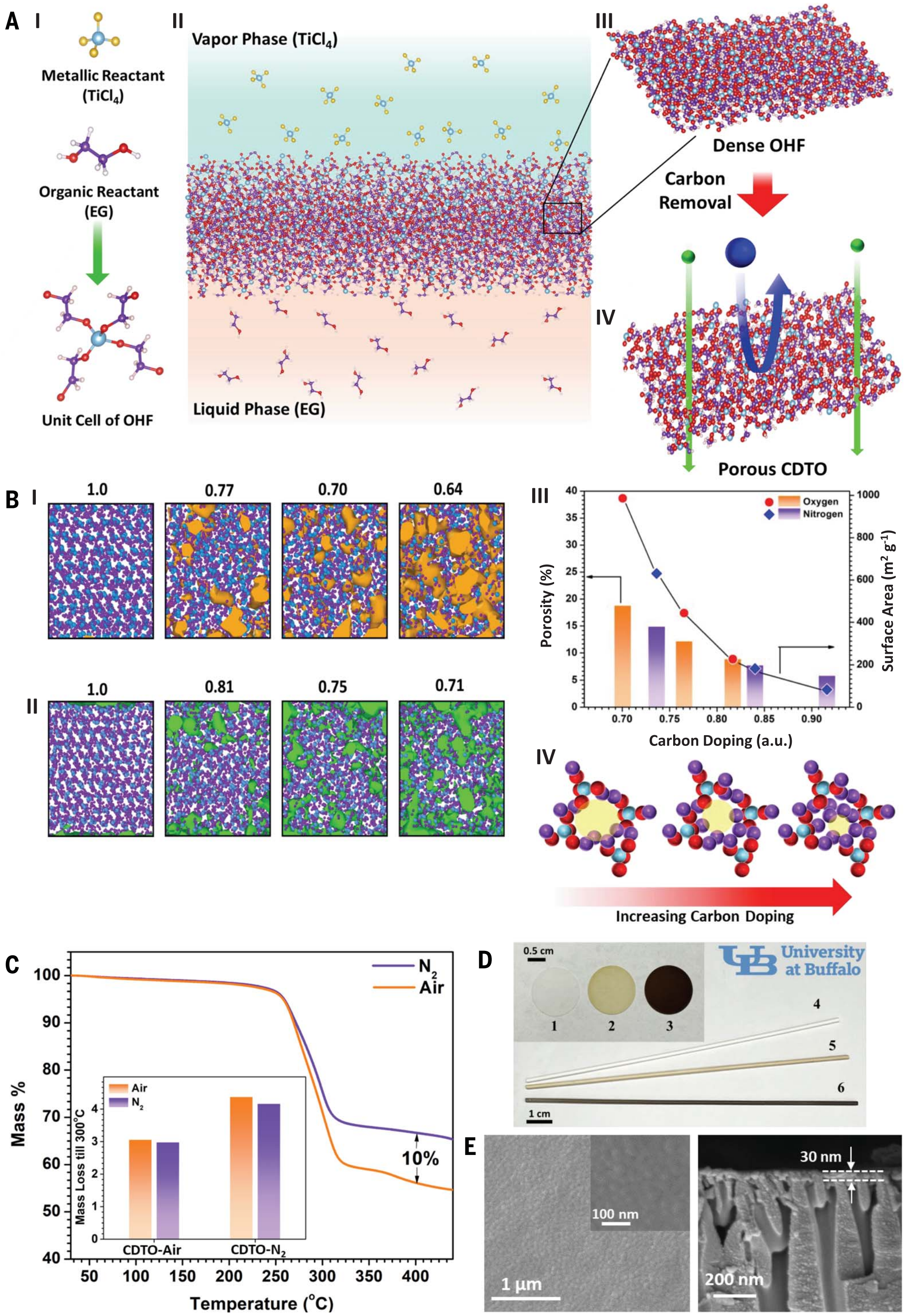


Fig. 1. CDTO nanofilm formation and characterization. (**A**) Schematic showing the formation of the OHF and subsequent generation of the porous CDTO through calcination. The reaction scheme (I) shows the formation of OHF through a facile substitution reaction between EG and $TiCl_4$ at the interface of two phases (II), leading to the formation of a dense OHF (III) that becomes porous CDTO (IV) after careful removal of carbon. (**B**) Evolution of nanopores in CDTO when OHF undergoes thermal treatment in O_2 (I) or N_2 (II), as realized by molecular dynamics simulation; blue and purple spheres represent titanium and carbon atoms, respectively, and solid parts are saffron or green and represent voids formed. (Left to right) Decrease in carbon doping (values indicated on top of each frame) and increase in pores/voids in the structure. CDTO's porosity and surface area (and hence pore size) can be precisely tuned over a large range by controlling the carbon doping (III and IV), as evidenced from the simulation. (**C**) Thermogravimetric analysis (TGA) of OHF showing its mass loss at 250°C to generate porous CDTO nanofilms. (Inset) mass loss of CDTO-Air and CDTO-N_2 in air and N_2, respectively, up to 300°C. (**D**) Pictures of OHF and CDTO nanofilms on anodic aluminum oxide (AAO) and α-alumina hollow fiber supports (1, 4: OHF; 2, 5: CDTO-Air; 3, 6: CDTO-N_2). (**E**) SEM images of the CDTO-Air nanofilm on AAO: surface (left, inset shows surface topography at higher magnification) and cross-section (right).

and precise size alteration (Fig. 1B, IV, and figs. S15 and S16). We speculate that carbon follows the fastest path to move from the bulk to gas phase, leaving behind voids that form highly dense, nanometer-sized pores, as observed in our simulation (fig. S12 and movies S1 and S2). OHF is stable (≤ 5% mass loss) in both air and N_2 until 250°C (Fig. 1C), beyond which it decomposes, possibly losing carbon as indicated in the simulation, with ~10% more mass loss in air than in N_2 at 400°C (fig. S17). However, porous CDTO is stable up to 300°C in both air and N_2 (Fig. 1C, inset), showing only 3 to 4% mass loss. Change in composition, as shown in x-ray photoelectron spectroscopy (XPS), and formation of oxygen-containing carbonaceous groups, as seen in Fourier transform infrared (FTIR) spectroscopy, also indicate carbon removal upon heating, consistent with our simulation results (figs. S18 and S19).

OHF was thus controllably calcined, either in air or N_2, at temperature ≥ 250°C for 2 hours, to precisely regulate the "carbon doping" to form porous CDTO nanofilms. We denominated CDTO-Air and CDTO-N_2 for OHF calcined at 250°C in air and N_2, respectively (unless stated). Table 1 summarizes elemental composition,

mechanical strength, and surface characteristics of the CDTO nanofilms. Although calcination makes CDTO porous, it maintains excellent mechanical stability with Young's modulus between 50 and 90 GPa (Table 1), comparable to the strongest reported OSN membrane (*10*). Calcination environment affects residual carbon; protective N_2 environment impedes carbon removal, evidenced by simulation and measured carbon content (Table 1, Fig. 1B, III). Controlling carbon content also allows surface property adjustment; greater hydrophobicity was realized with higher carbon doping (fig. S20). Hydrophobicity of OSN membranes is crucial as even a miniscule amount of water in a solvent can substantially foul hydrophilic membranes, limiting their industrial use (*4*). Figure 1D shows centimeter-scale OHF and CDTO nanofilms on AAO and HF with varying composition and properties; higher carbon doping (Table 1) leads to darker membrane color: yellow for CDTO-Air and black for CDTO-N_2. SEM images (Fig. 1E) show a selective layer of CDTO-Air, ~30-nm thick on AAO. Compared with the porous supports, the CDTO surface is smooth (Fig. 1E and figs. S21 and S22).

Nanopores are rigid, stable, and selective

We investigated transport of various organic solvents (table S3) through CDTO nanofilms and observed 2 to 3 orders of magnitude higher permeance compared with commercial OSN membranes, with hexane permeance of 550 L m^{-2} h^{-1} bar^{-1}. Interaction between solvents and CDTO was minimal, as suggested by weak dependence of viscosity-normalized permeance on Hansen solubility parameters (fig. S23), in agreement with other OSN membranes with rigid pores (*10*, *11*, *15*). Permeance is correlated to viscosity, as predicted by the Hagen-Poiseuille equation; a solvent—e.g., hexane—having the lowest viscosity, permeated fastest and vice versa, for both CDTO-Air and CDTO-N_2 nanofilms on AAO (Fig. 2A). Solvents permeated slower in CDTO-N_2 than CDTO-Air as a result of smaller pores generated during calcination in N_2 (higher carbon doping), consistent with our simulation (figs. S13 and S14).

Industrially, HF membranes are more applicable than flat sheets because of the higher module packing density and ability to withstand high pressure (*28*). Hence, we also prepared CDTO nanofilms on HFs and tested them for transport and separation efficiency (Fig. 2, B and C). Similar viscosity-dependent permeance was observed through CDTO nanofilms prepared on HFs (Fig. 2B). For dimethylformamide (DMF), a harsh solvent, when temperature was increased up to 140°C (causing viscosity decrease), the permeance followed the same viscosity correlation as other solvents at room temperature. This indicates exceptional rigidity of CDTO pores in harsh solvent, even at elevated temperatures. CDTO nanofilms on HFs, however, exhibited lower permeance than those on AAO. This difference in permeance is because of the formation of a thicker skin layer (~150 nm) on HFs, in contrast to ~30 nm on AAO, resulting from larger pore size of HFs and thus requiring longer time to form a dense OHF (Fig. 1E and figs. S7 and S22). Although the solvent permeance differs (fig. S24), the permeability (= permeance × thickness) is very similar (methanol: 6.6 and 6.3 L m^{-2} h^{-1} bar^{-1} μm for CDTO-N_2 on AAO and on HF, respectively), indicating that solvent transport depends only on CDTO material. Using a dead-end setup, a widely adopted technique for evaluating OSN membranes (*1*–*3*, *5*, *10*, *11*), we measured solute rejection (table S4) by CDTO-N_2 and CDTO-Air. Prior to conducting systematic solute rejection measurements using the dead-end system, we measured the rejection of a dye (Rose Bengal) using both the dead-end and crossflow systems. Our results showed similar rejection from both systems (fig. S25). This may suggest that under our testing conditions, the dead-end setup could provide reliable measurements for membrane performance evaluation. We found that the separation effectiveness of CDTO-Air and CDTO-N_2 prepared on both AAO and HF supports to be similar (Fig. 2C), confirming CDTO pores to be independent of the underlying supports.

For effective separation, pores in OSN membranes should be rigid in harsh solvents and at elevated temperatures. We observed stable separation of Rose Bengal from DMF by CDTO-Air nanofilm up to 140°C (Fig. 2D). This was also observed for other CDTO nanofilms and with different solvents (fig. S26). These nanofilms exhibited long term stability (over 3 months) in organic solvents with constant solute rejection, independent of solvents or operating conditions, predominantly through a "size-sieving" mechanism through its rigid molecular-sized pores (Supplementary Notes 2 and 3 and figs. S27 to S34). Other glycols can also form CDTO nanofilms (fig. S35).

Precise pore size control through carbon doping

For OSN, membranes with tunable nanopores with MWCO between 200 and 1400 g mol^{-1} are desirable, catering to diverse industrial processes. The residual carbon in the CDTO nanostructures dictates MWCO of these nanofilms, as suggested by simulation (Fig. 1B, III and IV, and fig. S16). We employed two strategies to vary carbon doping: (i) changing initial carbon content of the dense OHF and (ii) controlling carbon removal by altering calcination conditions (gas environment/temperature). This generated CDTO nanofilms with precisely controlled MWCO (Fig. 2E) covering the entire OSN range (*4*). Initial carbon content in OHF was controlled by addition of H_2O in EG (fig. S18). OHF formed by adding H_2O generated smaller pores in CDTO after calcination. This is evident from the decrease of MWCO from 620 to 240 g mol^{-1} for CDTO-N_2-H_2O (CDTO obtained by calcining OHF prepared by mixing H_2O in EG at 250°C) and from 920 to 320 g mol^{-1} for CDTO-Air-H_2O, as we increased water concentration (10 to 30 wt %) in EG (Fig. 2E and figs. S36 and S37). We speculate that water introduces Ti–O–Ti bonds, leading to less carbon (fig. S18) and thus generating smaller pores after calcination (fig. S38). Moreover, as calcination in N_2 generates smaller pores; CDTO-N_2-H_2O showed lower MWCO than did CDTO-Air-H_2O. We achieved the lowest MWCO of 240 g mol^{-1} by calcining OHF in N_2 prepared from EG with 30 wt % water. Hence, changing the initial carbon content of OHF allows MWCO tuning between 240 to 920 g mol^{-1} (Fig. 2E).

The other approach of tuning pore size is through controlling carbon removal (Fig. 2E and figs. S39 and S40). Calcining at higher temperatures or in air removes more carbon from OHF, hence reducing carbon doping (Table 1) and generating larger pores, as predicted by our simulation (Fig. 1B, III, and IV, figs. S13 to S16, and fig. S41). By increasing the calcination temperature from 250 to 500°C in air, we found that the MWCO of CDTO nanofilms increased from 920 g mol^{-1} to >1000 g mol^{-1} (Fig. 2E and fig. S39). Calcination at temperatures ≥ 300°C in air causes rapid carbon removal, with possible crystalline TiO_2 clusters (fig. S42), generating looser structures with larger pores and yielding membranes with

Table 1. Chemical composition, water contact angle, and Young's modulus of CDTO nanofilms and OHF. Temperature in membrane name represents the calcination temperature. Chemical composition is based on atomic percentage.

Membrane	Chemical composition (%)			Contact angle	Young's modulus
	Carbon	Titanium	Oxygen		GPa
OHF	58.1	11.5	30.4	93.4 ± 1.8°	61.5 ± 2.3
CDTO-Air (250°C)	36.3	19.3	44.4	42.0 ± 2.2°	55.1 ± 3.4
CDTO-N_2 (250°C)	39.7	17.4	42.9	78.0 ± 3.2°	76.4 ± 1.4
CDTO-Air (400°C)	23.5	21.9	54.6	30.9 ± 0.3°	50.6 ± 3.7
CDTO-N_2 (400°C)	38.6	20.3	41.1	75.3 ± 0.2°	88.8 ± 5.9

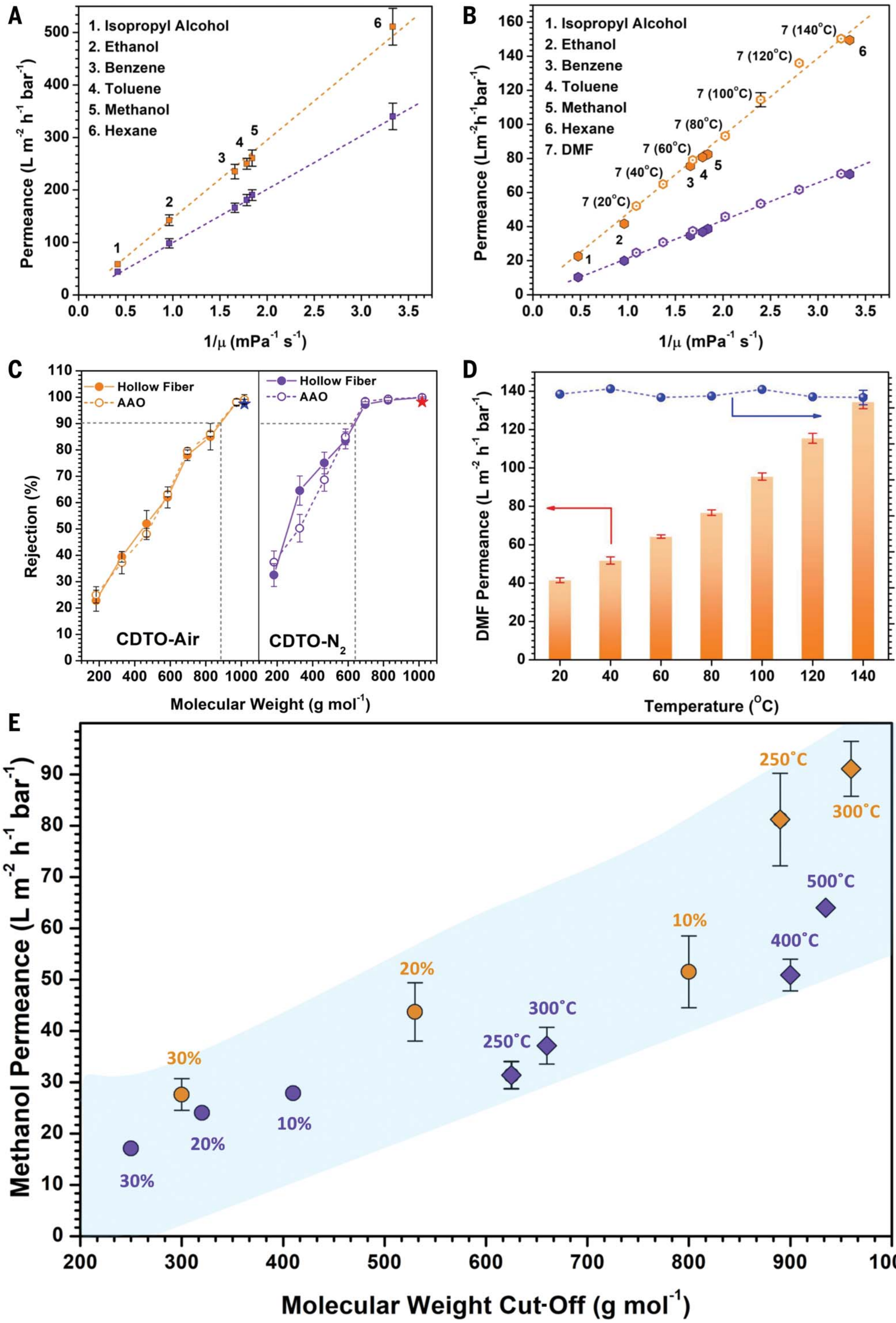


Fig. 2. Solvent permeation through and dye rejection by CDTO nanofilms. (**A**) Permeation of various solvents through CDTO nanofilms on AAO at 20°C. Orange represents CDTO prepared by calcination in air at 250°C, denoted as CDTO-Air, and violet represents CDTO prepared by calcination in N_2 at 250°C, denoted as CDTO-N_2. (**B**) Permeation of various solvents through CDTO nanofilms on α-alumina hollow fiber support (50-nm pores). Solid symbols for solvents 1 to 6 represent permeation at 20°C, and open symbols for solvent 7 represent permeation at 20 to 140°C. Orange represents CDTO-Air, and violet represents CDTO-N_2. (**C**) Dead-end dye rejection by CDTO nanofilms on AAO and α-alumina hollow fiber supports at 20°C: CDTO-Air (left); CDTO-N_2 (right). The stars represent rejection in a crossflow setup for HF membranes. (**D**) Separation performance of CDTO-Air on α-alumina hollow fiber for Rose Bengal solution in DMF at different temperatures. (**E**) Methanol permeance versus molecular weight cut-off (MWCO) at 20°C for CDTO nanofilms prepared on α-alumina hollow fiber support under different conditions. Violet represents CDTO calcined in N_2 and orange represents calcination in air. Diamond represents CDTO nanofilms prepared at different calcination temperatures (temperature values next to the symbols) and using pure ethylene glycol (EG). Circle represents CDTO nanofilms prepared by adding different weight percentages of water to EG (values next to the symbols) while maintaining the calcination temperature at 250°C. Error bars represent standard deviation from 3 samples. Error bars are omitted if they are smaller than symbols.

MWCO >1000 g mol^{-1}. Increasing calcination temperature from 250 to 500°C in N_2, MWCO for CDTO increased from 620 to 935 g mol^{-1} (Fig. 2E and fig. S40). Typically, membranes with lower MWCO have lower permeance and vice versa (Fig. 2E). CDTO thus demonstrates effective pore size tunability between 240 and 1400 g mol^{-1} (Fig. 2E and fig. S43, with MWCO step changes as small as 100 g mol^{-1}. This represents precise tunability, covering the broadest range for a single OSN membrane material (table S5) and allowing fabrication of OSN membranes with tailored pore sizes for specific applications within the entire OSN range.

High-density nanopores yields ultrafast transport

We believe densely packed nanopores with low tortuosity in CDTO are responsible for the ultrafast solvent transport. As observed, solvent transport through CDTO nanofilms is highly dependent on viscosity, with minimal interaction between solvent and the membrane material (fig. S23). In such a case, solvent transport can be described by the "pore-flow" model, following the Hagen-Poiseuille equation (*3*, *10*, *11*, *13*, *16*, *21*, *29*, *30*):

$$\text{Permeance} = \frac{J}{\Delta P} = \frac{\pi \varepsilon r_p^2}{8 \mu \delta \tau}$$

where J is solvent flux, ΔP is transmembrane pressure drop, ε is surface porosity, r_p is pore radius, μ is solvent viscosity, δ is membrane thickness, and τ is tortuosity. Although Hansen solubility correction may be required for some materials exhibiting considerable interaction with solvents, the nature of transport remains the same (*5*, *12*). This aligns with our observation of viscosity-dependent permeance (Fig. 2, A and B) and proportional increase of flux with transmembrane pressure drop (up to 35 bar, fig. S44). Although reducing "δ" proportionally increases permeance, high ε/τ—indicating high density of nanopores with low tortuosity—can also boost transport (Fig. 3A), eliminating challenging fabrication of membranes with "near atomic" thinness.

Properties of nanoporous membranes, including pore density, pore size and its distribution, and pore connectivity, are crucial for understanding and improving transport through membranes (*31*). The best estimation for these properties is made from measurements of permeation and separation performance of the nanoporous films (*32*). Thus, we calculated ε/τ of CDTO nanofilms and that of the reported OSN membranes, using pure methanol flux and the calculated effective pore size (calculation in Supplementary Note 4 and fig. S45), and compared them as a function of MWCO (range: 200 to 1400 g mol^{-1}). As shown in Fig. 3B, ε/τ of CDTO nanofilms increases with the increase of MWCO generally, with the highest value being 0.175. This general trend suggests that desired nanoporous structure (high ε/τ) for high permeation rate could be generated more easily for larger nanopores but could be more challenging for smaller nanopores. It is thermodynamically unfavorable to generate highly porous materials with smaller nanopores, as smaller pores tend to merge into larger ones, minimizing surface area and hence free energy. Materials that are relatively rigid and crystalline in nature may possess large amounts of small nanopores, such as MOFs, ZIFs, and zeolites. However, they usually lack the processibility of amorphous polymers and thus it is challenging to assemble them into thin, defect-free films.

Compared with commercial OSN membranes with the same MWCO, ε/τ of CDTO nanofilms is approximately 1 to 2 orders of magnitude higher. Compared with the OSN membranes reported in the literature, ε/τ of CDTO is 1.6 to 3 times higher in the MWCO range of 400 to 1000 g mol^{-1}. In the MWCO range of 200 to 300 g mol^{-1}, CDTO shows comparable ε/τ to the two best reported OSN membranes—diamond-like carbon (DLC) and 3D-COF. We speculate that during calcination, a large number of homogeneously distributed carbon atoms in OHF (Fig. 1B, I and II, and figs. S16 and S17) are removed from the structure, leaving behind evenly distributed, high-density small nanopores with low tortuosity and thus

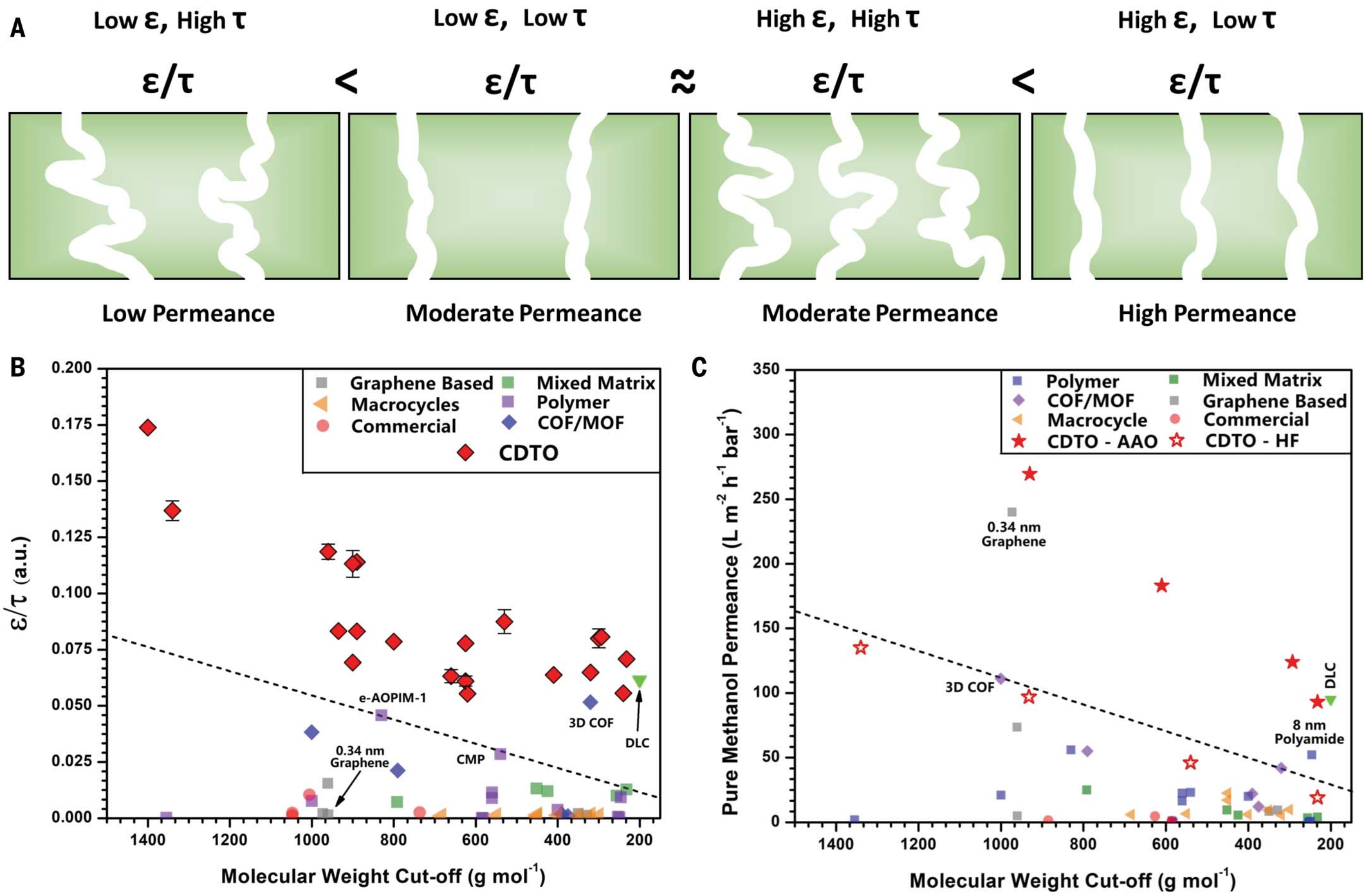


Fig. 3. Comparison of CDTO nanofilms with OSN membranes. (**A**) Schematic showing the influence of ε and τ on permeation through an OSN membrane when pore size and membrane thickness is fixed. For constant pore size and thickness, high density of pores having low tortuosity yields high permeance. (**B**) Comparison of ε/τ of CDTO membranes with commercial OSN membranes and state-of-the-art membranes reported in the literature. Only DLC (*10*) and 3D COF membrane (*74*) fall in the region of our CDTO membranes. (**C**) Separation performance comparison of our CDTO membranes with state-of-the-art commercial membranes and membranes reported in the literature. Pure methanol permeance at 20°C versus MWCO of these membranes is presented. CDTO nanofilms on AAO (high-permeance support) show the ultra-high permeance at a specific MWCO and are 2 to 3 orders of magnitude higher than commercial OSN membranes. CDTO nanofilms on α-alumina hollow fiber support (low-permeance support) have higher permeance than most OSN membranes, except for DLC (*10*), 8 nm polyamide (*5*) and 1 atom thin graphene (*14*). This is apparently because of the support resistance. The dashed black lines are only a guide for the eye.

conferring a high ε/τ (observed in our simulation, Fig. 1B). We believe that the high mechanical strength of the Ti-O network after carbon removal (Table 1) is essential to maintain a large number of small nanopores without them collapsing into larger ones. Coincidentally, the DLC membrane, which has ε/τ close to that of CDTO, also has a very high Young's modulus (*10*). The high ε/τ of CDTO nanofilms favors fast transport of solvent molecules and thus ensures their ultrahigh permeance, even if they are thicker or have comparable thickness to other OSN membranes (fig. S46). We compared our CDTO nanofilms' OSN performance with commercial and reported OSN membranes, in terms of MWCO and pure methanol permeance (Fig. 3C, Supplementary Note 5, and table S5). With the same MWCO, CDTO nanofilms on AAO (high permeance support) exhibit approximately 2 times higher pure methanol permeance than the highest reported values, apparently resulting from the high ε/τ. Even CDTO nanofilms on HF (low permeance support) are on par with the best reported membranes.

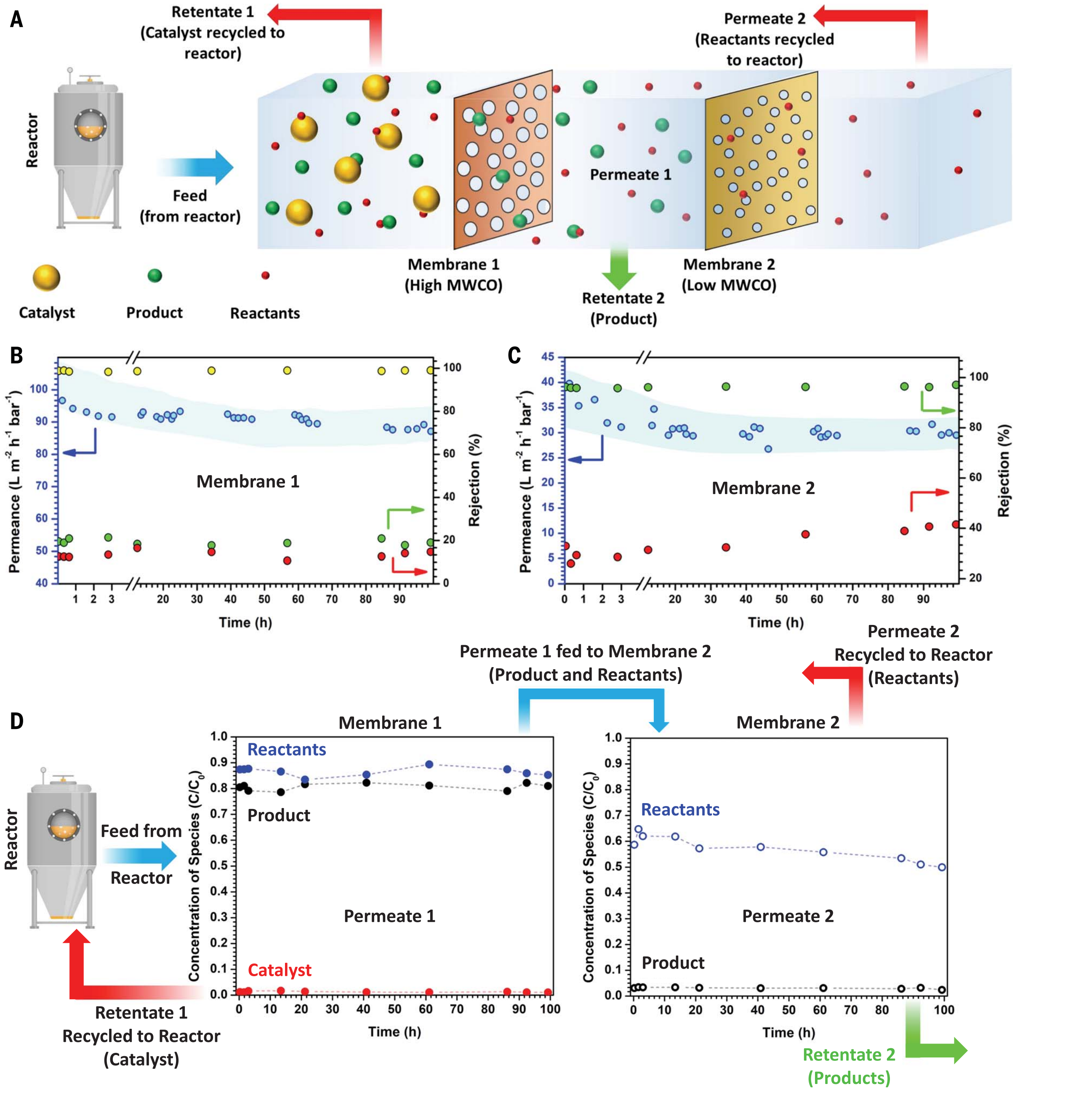


Fig. 4. Demonstration of CDTO membranes in the manufacturing of Boscalid under industrially relevant conditions. (**A**) A two-stage membrane cascade system designed using two CDTO membranes: membrane 1 with larger pores (MWCO: 940 g mol^{-1}) to retain the homogeneous catalyst (1150 g mol^{-1}); membrane 2 with smaller pores (MWCO: 300 g mol^{-1}) to separate product (340 g mol^{-1}) from reactants (~160 g mol^{-1}). Retentate from membrane 1 (containing catalyst) and permeate from membrane 2 (containing unreacted reactants) are to be recycled back into the reactor. (**B**) In 100-hours operation, membrane 1 shows >99% catalyst rejection and <10% permeance drop. (**C**) In a 100-hour operation, membrane 2 shows >95% retention of product while allowing reactants to pass through. (**D**) Concentrations of catalyst, product, and reactants over time for permeates 1 and 2 [as defined in (A)]. The concentrations of the components in each stream have been normalized by their concentration in the feed.

Industrial separation under harsh conditions

Specialty chemicals such as small drugs, agrichemicals, and others can require challenging synthesis conditions that involve harsh solvents at high temperatures and pressures. Membranes that are stable under such conditions can significantly reduce energy consumption by separating products and catalysts from reactors at the synthesis conditions. Whereas most polymeric membranes may fail, including commercial OSN membranes (fig. S47), CDTO membranes are stable and expected to be effective for such applications. To demonstrate this, we selected CDTO membranes with appropriate MWCOs and used them to separate reactants, product, and homogeneous catalyst in the production of the pesticide Boscalid (*33*) under industrially relevant conditions. We designed a two-stage membrane cascade system using two appropriate CDTO membranes with MWCOs of 940 and 300 g mol^{-1} respectively, to separate (i) catalyst from reactants and product and (ii) product from reactants at 90°C in DMF (Fig. 4A and figs. S48 to S51). Preliminary dead-end screening indicated the membranes to be effective for the separation (table S6). Continuous crossflow operation for 100 hours demonstrated CDTO's capability for sustained rejection of catalyst and product (consistent with the dead-end screening) with a separation factor of 65.9 (product/catalyst) and 17.4 (reactants/product) for the looser membrane 1 (higher MWCO) and tighter membrane 2 (lower MWCO), respectively, while being stable in DMF at 90°C, close to actual synthesis conditions (Fig. 4, B and C). These high separation factors (comparison with literature in fig. S52) led to highly effective catalyst recovery (<1% loss) and extraction of 80 to 90% of the reactants/product by membrane 1 and highly efficient reactant recovery (with <5% product) for recycling.

Conclusion

In summary, our work addresses the need for inorganic counterparts to interfacial polymer membranes, which can be fabricated into defect-free nanofilms through fast interfacial reaction. CDTO nanofilms are stable in harsh solvents at temperatures up to 140°C and have rigid nanopores that can be precisely controlled within the entire OSN range, exhibiting a broad and precise pore tunability for a single OSN membrane material. CDTO nanofilms exhibit a high ε/τ, probably resulting from their high mechanical strength that enables high-density, evenly distributed nanopores. This allows CDTO nanofilms to exhibit high permeance, even if they are not atomically thin. Owing to their stability, these membranes may have applications in industrial processes involving harsh conditions. Additionally, other suitable metallic and organic reactants may also be explored in the future to fabricate such interfacially generated nanofilms for separation applications.

REFERENCES AND NOTES

1. Z. Jiang *et al.*, *Nature* **609**, 58–64 (2022).
2. S. Li *et al.*, *Science* **377**, 1555–1561 (2022).
3. L. Huang *et al.*, *Adv. Mater.* **28**, 8669–8674 (2016).
4. P. Marchetti, M. F. Jimenez Solomon, G. Szekely, A. G. Livingston, *Chem. Rev.* **114**, 10735–10806 (2014).
5. S. Karan, Z. Jiang, A. G. Livingston, *Science* **348**, 1347–1351 (2015).
6. M. F. Jimenez-Solomon, Q. Song, K. E. Jelfs, M. Munoz-Ibanez, A. G. Livingston, *Nat. Mater.* **15**, 760–767 (2016).
7. B. Liang *et al.*, *Nat. Chem.* **10**, 961–967 (2018).
8. K. A. Thompson *et al.*, *Science* **369**, 310–315 (2020).
9. E. J. Kappert *et al.*, *J. Membr. Sci.* **569**, 177–199 (2019).
10. S. Karan, S. Samitsu, X. Peng, K. Kurashima, I. Ichinose, *Science* **335**, 444–447 (2012).
11. Q. Yang *et al.*, *Nat. Mater.* **16**, 1198–1202 (2017).
12. C. Cheng, S. A. Iyengar, R. Karnik, *Nat. Nanotechnol.* **16**, 989–995 (2021).
13. L. Nie *et al.*, *Sci. Adv.* **6**, eaaz9184 (2020).
14. L. Shen *et al.*, *Sci. Adv.* **7**, eabg6263 (2021).
15. D. B. Shinde *et al.*, *Chem. Sci.* **11**, 5434–5440 (2020).
16. D. B. Shinde *et al.*, *J. Am. Chem. Soc.* **140**, 14342–14349 (2018).
17. M. Amirilargani *et al.*, *J. Membr. Sci.* **564**, 259–266 (2018).
18. T. Liu *et al.*, *Mater. Chem. Front.* **5**, 3184–3191 (2021).
19. H. Li *et al.*, *Science* **367**, 667–671 (2020).
20. T. E. Culp *et al.*, *Science* **371**, 72–75 (2021).
21. T. Huang *et al.*, *Nat. Commun.* **11**, 5882 (2020).
22. T. Huang, T. Puspasari, S. P. Nunes, K. Peinemann, *Adv. Funct. Mater.* **30**, 1906797 (2020).
23. L. Wang *et al.*, *Nat. Nanotechnol.* **12**, 509–522 (2017).
24. B. Mi, *Science* **364**, 1033–1034 (2019).
25. X. Meng, *J. Mater. Chem. A.* **5**, 18326–18378 (2017).
26. Z. Song *et al.*, *J. Membr. Sci.* **510**, 72–78 (2016).
27. A. P. Thompson *et al.*, *Comput. Phys. Commun.* **271**, 108171 (2022).
28. P. Yazgan-Birgi, M. I. H. Ali, H. A. Arafat, *Separ. Purif. Tech.* **212**, 709–722 (2019).
29. L. Zhang, M. Zhang, G. Liu, W. Jin, X. Li, *Adv. Funct. Mater.* **31**, 2100110 (2021).
30. X. He *et al.*, *Adv. Funct. Mater.* **29**, 1900134 (2019).
31. H. W. H. Lai *et al.*, *Science* **375**, 1390–1392 (2022).
32. B. Sutariya, S. Karan, *Sep. Purif. Technol.* **293**, 121096 (2022).
33. R. Raghavan, J. B. Joshi, *Chem. Eng. Prog.* **114**, 49–56 (2018).

ACKNOWLEDGMENTS

Funding: B.S. and M.Y. acknowledge National Science Foundation (Fund No: 1451887); Q.B. and D.B. acknowledge the University at Buffalo Start-up Fund; M.Y. acknowledges Department of Energy (Fund No: DE-FE-0031730) and G.B. acknowledges the Department of Energy, Basic Energy Sciences Division (DE-FG02-09ER16005). **Author contributions:** M.Y., B.S., and Q.D. conceived the project. Q.D. and B.S. carried out experiments for membrane fabrication and testing. B.S. performed contact angle, FTIR, XRD, XPS, EDS, and TGA tests. Q.D. performed some XPS and XRD tests. Q.D. and J.J. collected SEM images. R.K. and P.K. performed the MD simulations. R.Y. and J.L. tested mechanical properties. B.S., M.Y., D.B., and G.B. evaluated and discussed data. M.Y. directed all aspects of the project. B.S. and M.Y. wrote the manuscript. All authors reviewed and discussed the manuscript. **Competing interests:** A US provisional patent (application number: PCT/US2023/022354), with M.Y., B.S., and QD as co-inventors, has been filed by U.B. **Data and materials availability:** All data are available in the main text or the supplementary materials.

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh2404
Materials and Methods
Supplementary Text
Figs. S1 to S52
Tables S1 to S6
References (*34*–*76*)
Movies S1 and S2

Submitted 18 February 2023; accepted 11 August 2023
10.1126/science.adh2404

BIOMEDICINE

Implantable bioelectronic systems for early detection of kidney transplant rejection

Surabhi R. Madhvapathy[1,2], Jiao-Jing Wang[3], Heling Wang[4,5,6], Manish Patel[2,7], Anthony Chang[8], Xin Zheng[3], Yonggang Huang[4,5], Zheng J. Zhang[3,9,10]*, Lorenzo Gallon[3,11]*†, John A. Rogers[1,2,12,13]*†

Early-stage organ transplant rejection can be difficult to detect. Percutaneous biopsies occur infrequently and are risky, and measuring biomarker levels in blood can lead to false-negative and -positive outcomes. We developed an implantable bioelectronic system capable of continuous, real-time, long-term monitoring of the local temperature and thermal conductivity of a kidney for detecting inflammatory processes associated with graft rejection, as demonstrated in rat models. The system detects ultradian rhythms, disruption of the circadian cycle, and/or a rise in kidney temperature. These provide warning signs of acute kidney transplant rejection that precede changes in blood serum creatinine/urea nitrogen by 2 to 3 weeks and approximately 3 days for cases of discontinued and absent administration of immunosuppressive therapy, respectively.

Approximately 60% of the 40,000 solid organ transplants performed in the US in 2022 were kidneys (*1*). However, human leukocyte antigen genotype mismatch between the donor and recipient can result in transplant rejection (*2*). Graft failure can occur at any time: 1-year, 5-year, and 10-year graft survival rates are 92.7, 77.6, and 49.5%, respectively, with rejection as the main cause of failure (*3*). Subclinical rejections can occur in 10 to 15% of renal transplant recipients within the first few months to a year post transplant (*4*–*7*).

Therapeutic intervention upon detection of early-stage rejection could preserve graft function. The current "gold standard" for detecting transplant rejection is a biopsy of the kidney cortex tissue, which has potential for complications such as bleeding, pain, infection, and accidental damage to adjacent organs (*8*). Biopsies are thus infrequent, typically performed 1 to 2 times in the first 24 months after transplantation or upon detection of elevated levels of serum creatinine/blood urea nitrogen (BUN) (*4*, *5*, *9*). Factors unrelated to rejection, such as diet, muscle mass, infection, and medications also alter these biomarkers, leading to false-negative or -positive predictions of rejection (*10*). Moreover, changes in serum creatinine lag from days to weeks behind changes in glomerular filtration rate (GFR) (*10*). Other minimally invasive techniques offer significant advantages in sensitivity but are limited by drawbacks similar to those in blood collection such as restricted monitoring frequencies and the requirement for off-site analysis (*11*–*18*). There remains a clinical need for techniques that can continuously monitor graft health from the moment of transplantation and detect the onset or early stages of rejection, when serum creatinine/BUN are still within the normal physiological range (subclinical rejection).

During acute rejection, inflammatory events mediated by T-cell–generated cytokines and/or by allo-antibodies binding to tissue antigens take place within the transplanted organ. We report fully implantable, wireless bioelectronic systems capable of precise measurements of local organ temperature (T_{kidney}) and thermal conductivity (k_{kidney}) as biophysical surrogates for inflammation/perfusion in rat renal transplant models. The sensor interfaces directly with the curved, soft surface of the organ for continuous in vivo monitoring in freely moving animals from the moment of transplantation.

[1]Department of Materials Science and Engineering, Northwestern University, Evanston, IL, USA 60208. [2]Querrey Simpson Institute for Bioelectronics, Northwestern University, Evanston, IL, USA 60208. [3]Comprehensive Transplant Center, Northwestern University, Chicago, IL, USA 60611. [4]Department of Mechanical Engineering, Northwestern University, Evanston, IL, USA 60208. [5]Department of Civil Engineering, Northwestern University, Evanston, IL, USA 60208. [6]Laboratory of Flexible Electronics Technology, Tsinghua University, Beijing, 100085 China. [7]Department of Intervention Radiology, University of Illinois at Chicago, Chicago, IL, USA 60612. [8]Department of Pathology, University of Chicago, Chicago, IL USA 60637. [9]Department of Surgery, Northwestern University Feinberg School of Medicine, Chicago, IL, USA 60611. [10]Simpson Querrey Institute for BioNanotechnology, Northwestern University, Chicago, IL, USA 60611. [11]Department of Nephrology, Northwestern University, Chicago, IL, USA 60611. [12]Department of Biomedical Engineering, Northwestern University, Evanston, IL, USA 60208. [13]Department of Neurological Surgery, Northwestern University, Chicago, IL, USA 60611.
***Corresponding author. Email: zjzhang@northwestern.edu (Z.J.Z.); l-gallon@northwestern.edu (L.G.); jrogers@northwestern.edu (J.A.R.)**
†These authors contributed equally to this work.

Monitoring rat kidney transplant rejection

Kidney transplant models based on rats are suited for research because they are well-characterized, highly repeatable, and cost effective. The surgical procedures used in the following studies involve removing both native kidneys and grafting the transplanted kidney through the vascular anastomoses between the donor vessels and the abdominal aorta/inferior vena cava on the right side of the body (Fig. 1A). The bioelectronic system, which is implanted after reperfusion of the grafted kidney, lies completely within the abdominal cavity and consists of a microfabricated, mechanically compliant biosensor that directly mounts on the surface of the kidney and connects by fine wires to a wireless electronics module secured to the adjacent abdominal wall. Additional images of the device and implantation appear in figs. S1 and S2.

The rat kidney is small (~1 × 1 × 2 cm^3), soft [Young's modulus (Y) ~4.5 kPa] (*19*), and highly perfused [~300 ± 51.3 ml × min^{-1} × (100 g tissue)$^{-1}$] (*20*). The miniature (~0.3 × 0.7 cm^2), stretchable (20% stretchability, fig. S3A), flexible (bending radius ≥ ~2.8 mm, fig. S3B), ultrathin (~220 μm), and smooth (0.13 μm ± 0.02 areal surface roughness) design of the sensor allows for a gentle and seamless interface to the delicate surface of the kidney without risk of organ damage. The biosensor relies on lithographically patterned features of gold (100-nm thick, ~20-μm wide) to form a 3-mm diameter thermal sensing "disk" and a pair of serpentine interconnects (each ~80-μm wide), all encapsulated by layers of polyimide (10 μm) and silicone (100 μm) (figs. S4 and S5). The sensor contacts the dorsal kidney cortex beneath a tight "pocket" formed under the ~25 μm thick renal capsule (*21*) and is sutured to it [Fig. 1, A (inset) and B]. The electronics module contains sensor readout circuitry and a bluetooth low energy system-on-chip for control and data transmission, with a coin-cell battery for power supply (fig. S6). The small lateral dimensions (~1.1 × 1.6 cm^2), thickness (~0.5 cm), and smooth surface of the electronics module facilitate insertion within the abdominal cavity (Fig. 1C). Details on the measurement appear in the methods section and figs. S7 to S9. Real-time, continuous data collection in untethered, freely moving animals, with stable operation is possible for 2 months or more (fig. S10) without adverse effects.

Measurements of T_{kidney} and k_{kidney} performed using this system for t = 27 days in the native (right) solitary kidney of a control animal highlight effects of surgical recovery and natural biological variations (Fig. 1D). T_{kidney} increases rapidly after the anesthesia wears off (t = 0 days) to a peak value ~3 to 4 hours post surgery (~39°C) and then subsequently decreases, concurrent with the release of post-operative analgesia to an average value of ~37.5°C. The irregular pattern in T_{kidney} between t = 0 to 2 days arises from post-operative recovery and from the influence of a heating pad under one side of the animal cage to assist with body temperature regulation. Details of the effects of this pad and of the ambient temperature are in fig. S11; variations in the ambient temperature do not influence T_{kidney}. A periodic ~1-day circadian rhythm emerges

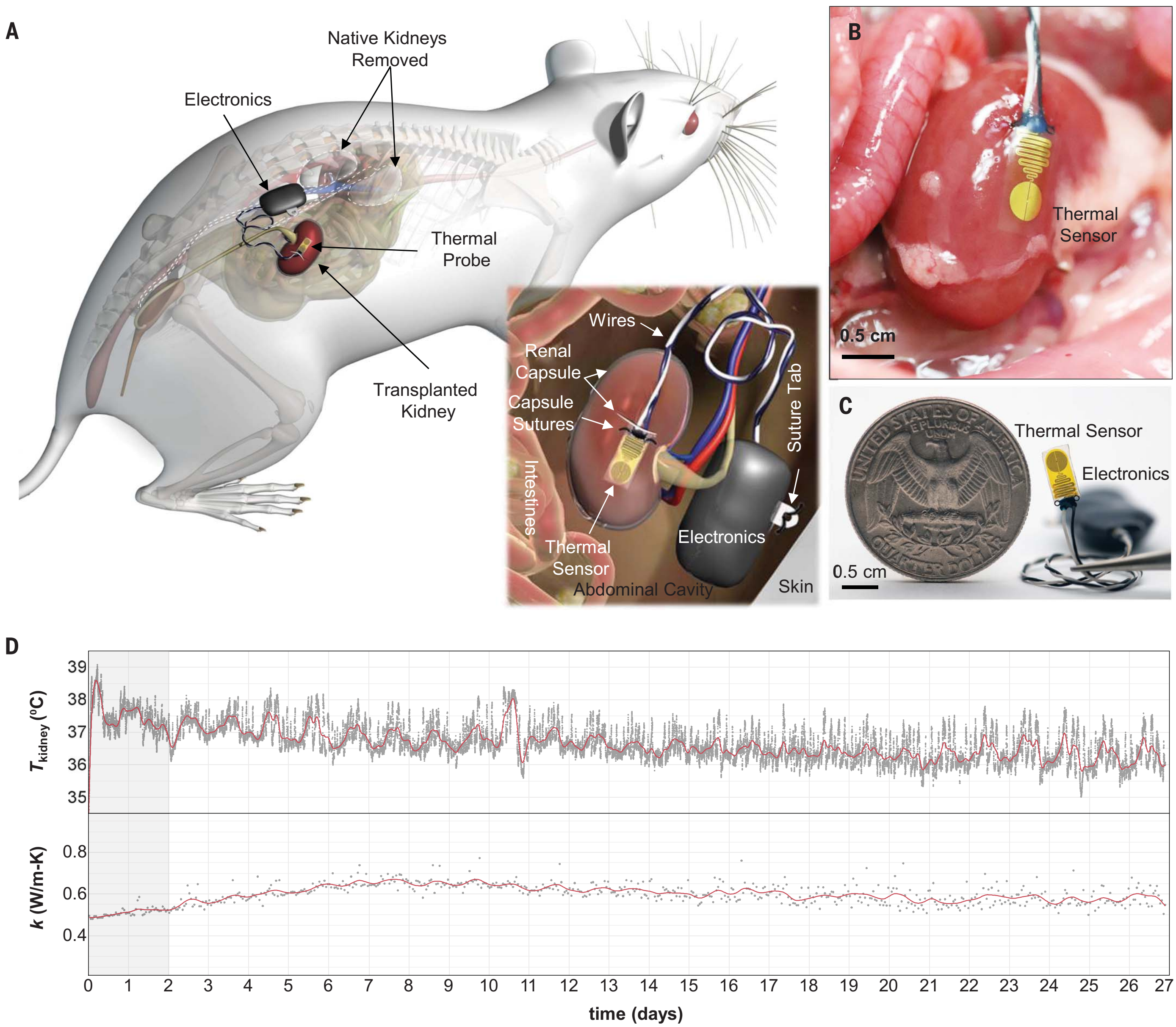


Fig. 1. Monitoring kidney transplant rejection using an implantable bioelectronic system. (**A**) Illustration of kidney transplant and device implantation in a rat model. The dashed white lines highlight removal of both native kidneys. The sensor directly interfaces with the cortex, sutured to the overlying renal capsule through two suture holes (inset). The electronics reside adjacent to the kidney, secured to the abdominal wall through a suture tab. The wires connecting the sensor and electronics module lie against the abdominal fat. Photograph of (**B**) the sensor implanted on the dorsal side of the kidney and (**C**) the device next to a US Quarter (scale bar = 5 mm). (**D**) Time variation of the organ temperature (T_{kidney}) and thermal conductivity (k_{kidney}) collected for a Lewis rat with a single native kidney (second kidney nephrectomized) and for t = 0 to 28 days. The gray points represent the raw data. The red line is a smoothing spline fit (λ = 0.01).

after t = 2 days, consistent with the return to a healthy behavioral pattern. The bioelectronic system has no measurable effect on grooming, activity, food and water consumption, or sleep/wake cycles. Short-term (~40 min to 1 hour) variations in T_{kidney} are correlated with motion/activity (fig. S12).

k_{kidney} is a function of perfusion and the tissue thermal properties, increasing from ~0.48 W/m-K to a final value of ~0.64 W/m-K over t = 0 to 6 days. k_{kidney} at t = 6 days for an animal with both native kidneys intact (fig. S13) is approximately half (~0.33 W/m-K) that of an animal with a single native kidney, consistent with the understanding that the vascular load of the body splits equally between the two kidneys. These observations in control animals provide a point of reference for the studies on transplant rejection.

Characterization of acute kidney transplant rejection

Renal transplantation between inbred rat strains with different major histocompatibility complexes (MHC) provides control over transplant acceptance or rejection. Lewis Rats (MHC haplotype $RT1^{l}$) are both the donors and recipients in isogeneic transplants (Fig. 2A). Isogeneic transplants are analogous to those between identical twins in humans and result in graft acceptance without the need for immunosuppression. August-Copenhagen-Irish (ACI) rats (MHC haplotype $RT1^{av1}$) are the donors and Lewis rats are the recipients in allogeneic transplants (Fig. 2B). Allotransplants, which represent the

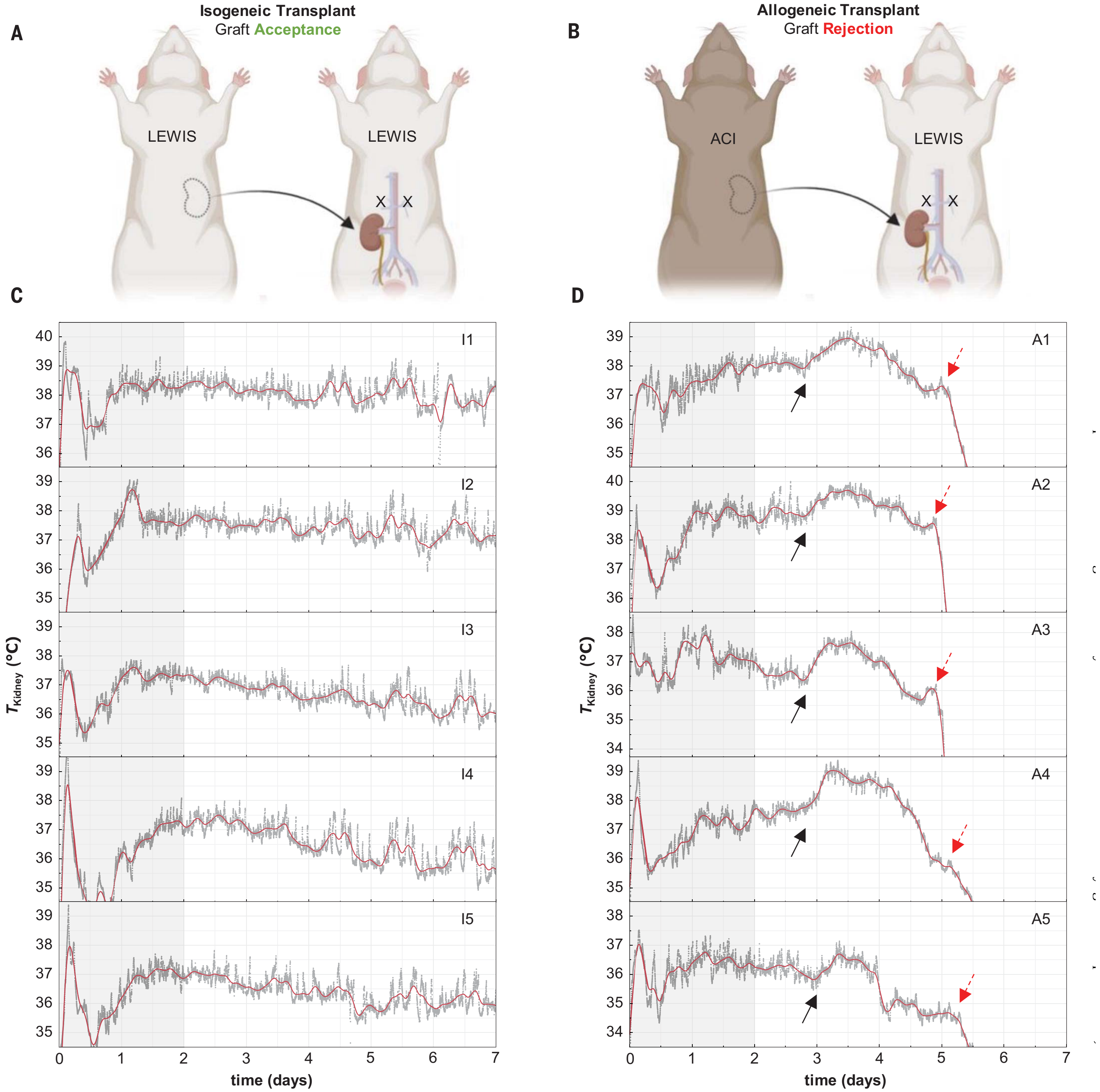


Fig. 2. Characterization of acute rejection. Illustrations of a rat kidney transplantation model using inbred rat strains for (**A**) isogeneic transplant, where both the donor and recipient strains are Lewis rats, leading to graft acceptance and (**B**) allogeneic transplant, where the donor strain is the August Copenhagen Irish (ACI) rat and the recipient is a Lewis rat, resulting in graft rejection. "X" denotes the removal of both recipient native kidneys. T_{kidney} measured for ~7 days for (**C**) $n = 5$ isografts and (**D**) $n = 5$ allografts. In this and subsequent figures, labels (e.g., A1-5, I1-5) identify each individual dataset (table S1). The gray points represent the raw data. The red line is a smoothing spline fit (λ = 0.01). The shaded region (t = 0 to 2 days) indicates the post-surgery recovery period during which a heating pad placed underneath half of the animal cage assists with thermoregulation. The black arrows in Fig. 2D correspond to a feature in T_{kidney} (bump and inflection point) unique to the allografts at t ~ 3 days. The red arrows in Fig. 2D correspond to a sharp decrease in T_{kidney} at t ~ 5 days.

most clinically relevant cases for humans, result in graft rejection. The survival time for ACI-to-Lewis transplants is ~8 days without immunosuppressive medication (*22*). Details of the procedures are in the methods section.

T_{kidney} for isografts varies until t ~ 3 days as a result of induced inflammation, effects of analgesia, and post-operative care (Fig. 2C). A circadian rhythm emerges after t ~ 3 days. The average daily T_{kidney} remains constant after t ~ 7 days (data >7 days is in fig. S14). Trends in T_{kidney} and overall behaviors are consistent for all $n = 5$ animals, similar to those of control animals with native kidney(s) (Fig. 1E and fig. S13). Minimal adhesions/foreign body response

(FBR) appear on the surface of the kidney or around the sensor, interconnecting wires, or electronics module at the experimental end point (21 to 28 days) (fig. S15). The response time (t_{90}) of the biosensor for measurements of temperature is 0.13 s (fig. S16). Growth of FBR on chronic timescales (*23*) increases this response time (t_{90} = 1.49 s for a fibrotic capsule thickness of 250 μm) but does not affect measurement sensitivity to T_{kidney} (fig. S17). These observations suggest that the graft is healthy, without adverse effects induced by the bioelectronic system.

T_{kidney} for allografts bears little resemblance to the data for isografts (Fig. 2D). For all allografts (n = 5), T_{kidney} rises at t ~ 3 days for a period of ~18 hours and then decreases over the subsequent ~18 hours. T_{kidney} decreases sharply (approximately −0.5°C per hour) around t ~ 5 to 6 days and 30 to 32°C, which establishes the experimental end point (full temperature range appears in fig. S18). Lack of food/water intake and locomotion characterize behaviors during the T_{kidney} decline on day 5/6 (fig. S12). At the end point, adhesions and FBR surround the graft, which is enlarged (~1.5 to 2 times the original size) and has a marbled appearance with necrotic patches distinctive of acute rejection (fig. S15). The impact of patchy or focal rejection on T_{kidney} appears in fig. S19.

The results for k_{kidney} for all isografts and allografts (fig. S20) are similar to those for a single healthy kidney (~0.64 W/m-K), as observed in fig. S13. In addition, k_{kidney} decreases with time for 3/5 allotransplants, consistent with the severe FBR observed at the end point for these cases. Unlike T_{kidney}, k_{kidney} does not display systematic trends that distinguish allografts from isografts. A drop in GFR may be obscured by any comparatively small change in perfusion/associated kidney tissue damage due to the rejection response. Second, the speckled/irregular color of rejected kidneys suggests that changes in perfusion, tissue necrosis, or other effects that can alter k_{kidney} may be localized and nonuniform over the kidney surface. In such cases, the measurements would depend on the relative size and exact position of the biosensor, in ways that are difficult to control systematically.

Organ temperature provides an early warning sign of transplant rejection

Correlating T_{kidney} with histological and blood biomarker analysis is essential for evaluating its significance and utility for detecting graft rejection. The following discussion (Fig. 3) focuses on two separate time points coinciding with the allograft T_{kidney} features identified in Fig. 2D: (i) t = 5 to 6 days (referred to as the end point) corresponding to the steep temperature decline and (ii) t = 3 to 4 days (referred to as the midpoint) corresponding to the inflection point and transitory rise and fall (i.e., "bump"). Isograft data serve as the control reference. Macroscopic and microscopic histological examination at the end point reveals that the kidney is normal for isografts and acutely rejected for allografts (Fig. 3, A and B, and extended data in figs. S21 and S22). Diffuse cortical necrosis and thrombotic microangiopathy—characteristic of severe acute rejection—appear in 4 of the 5 allografts. The fifth allograft exhibits type I tubulointerstitial rejection. BUN and creatinine levels are within the normal range for Lewis rats in the case of isografts (Fig. 3, C and D) but are highly elevated for allografts (P < 0.0001 and P < 0.0003, respectively). The high value of BUN (>140 mg/dl) for allografts indicates uremia, consistent with behavioral characteristics including signs of pain, slowed or absent motion, and reduced food and water intake. $\overline{T_n}$ denotes the average T_{kidney} between day n and n − 1. The temperature decline for the allografts, denoted by the difference in average T_{kidney} on the final and penultimate days of survival ($\overline{T_6}$ − $\overline{T_5}$), is also significant (P = 0.0122) (Fig. 3E). Thus, behavioral, blood marker, and T_{kidney} analysis accurately detect late-stage graft rejection at the end point, as verified by histology.

Evaluations of the data at the midpoint offer physiological insight into the origin of the T_{kidney} inflection point and bump seen in Fig. 2D. A separate set of experiments involves harvesting allograft and isograft kidneys close to the midpoint for this purpose (T_{kidney} and k_{kidney} data in fig. S23). For n = 3 cases each, histological examination reveals that the isografts have normal morphology whereas all allografts exhibit type I tubulointerstitial rejection (representative cases in Fig. 3, F and G; remaining data shown in fig. S24). The allograft case presented in Fig. 3G, harvested exactly at the inflection point (t ~ 3 days), displays mild type I tubulointerstitial rejection, demonstrating that the allograft T_{kidney} bump coincides with the histological onset of rejection.

By comparison, BUN and serum creatinine levels at the midpoint (t = 4 days) do not offer clear diagnostic value, as the difference in values for isografts (n = 4) and allografts (n = 6) is not statistically significant (Fig. 3, H and I). By contrast, ($\overline{T_4}$ − $\overline{T_3}$), which is an indicator of the height of the bump, is larger for allografts (~0.6°C) compared to isografts (approximately −0.3°C) (P = 0.0016, n = 5) (Fig. 3J). The bump therefore identifies acute rejection at an earlier time point than does BUN or serum creatinine. The absolute height and width of the allograft bump (n = 5) are ~1.0 ± 0.2°C and ~1.8 ± 0.4 days, respectively. The onset of the bump occurs at 2.8 ± 0.1 days. Animal behavior for allografts at t ~ 4 days appears normal and is indistinguishable from isografts. Thus, continuous measurements of T_{kidney} show potential for detection of acute rejection before signs of loss of graft function appear in blood markers or behavioral patterns. These features in T_{kidney} are unlike those observed in body T of rodent models of induced septic shock (*24*–*27*) or in T_{kidney} during renal ischemia reperfusion injury (fig. S25 and table S2).

Delayed graft rejection with immunosuppressants

Real-world kidney allotransplants require immunosuppressants to inhibit graft rejection. The following experiments (Fig. 4) mimic clinical cases of patient noncompliance, where patients prematurely stop or reduce their prescribed immunosuppressant dose. Allografts receive 1 mg/kg per day FK506 for t = 0 to 7 days through a subcutaneously implanted osmotic pump (2ML1, Alzet, Inc.) (Fig. 4A). The recordings include both T_{kidney} and k_{kidney} for one case, and measurements of only T_{kidney} for the other cases, using a simplified version of the device placed in contact with the kidney (fig. S26 showcases correlation for the gold versus contact sensor). This simplified device has extended battery lifetime (details in Methods). The time evolution of renal function (collected at fixed times t ~ 4, 7, 10, 14, 21, and 27 days) illustrates that creatinine and BUN rise above normal levels only at t ≥ 27 days (Fig. 4B), coincident with a ~9% decline in body weight (fig. S27).

T_{kidney} from isografts serves as a point of reference for medicated allograft data (Fig 4C). Data from allografts treated with 1 mg/kg FK506 exhibit several notable features that are unlike those observed in isografts (Fig 4, D to H). After the initial t = 0 to 2 day surgical recovery phase, T_{kidney} remains flat until t ~ 7 days, with minimal variations and no circadian cycle, concurrent with the administration period of FK506. An inflection point at t ~ 8 to 9 days appears in 4 of the 5 animals (Fig 4, D to G). T_{kidney} slowly rises and reaches a peak at t = 14 days, followed by a slow decline (a "bump"). At t ~ 22 days, T_{kidney} falls steeply (approximately −0.83°C per day on average). Data from medicated allograft cases also present additional higher frequency components relative to isografts. Data collected beyond t = 28 days, up to 2.5 months, appears in fig. S28.

Histopathology offers insight into kidney health at the time points corresponding to the occurrence of key features in T_{kidney} in Fig. 4, D to H ("bump", high-frequency components). Extended histology data and T_{kidney} for experiments that involve harvesting the kidney at earlier time points are in figs. S29 and S30. The kidney at t ~ 10 days is not rejected (Fig. 4I). At the peak in temperature (t ~ 14 days), and at t = 21 and 27 days, type I acute tubulointerstitial rejection occurs, characterized by

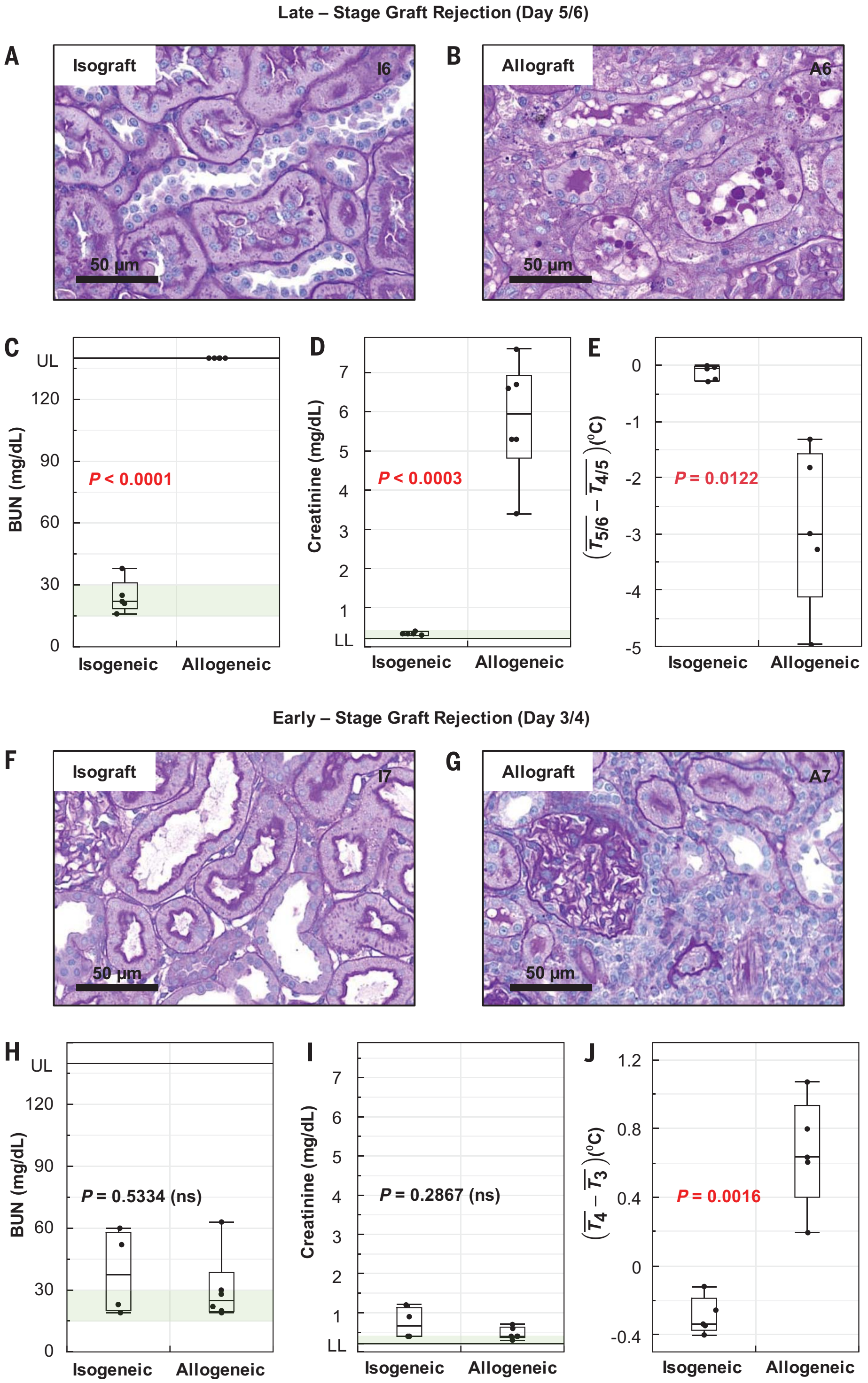


Fig. 3. Kidney temperature as an early indicator of acute rejection. Representative periodic acid-schiff (PAS)–stained histological section for an (**A**) isograft at t = 6 days, displaying normal parenchyma with back-to-back arrangement of tubules, normal glomeruli, and an absence of interstitial inflammation and (**B**) allograft at t = 6 days, with severe type I rejection characterized by prominent interstitial inflammation with tubulitis. Allografts show elevated (**C**) BUN and (**D**) serum creatinine levels at t ~ 6 days relative to isografts. UL and LL denote the upper and lower detection limits of the analyzer. $\overline{T}_n$ represents T_{kidney} averaged over $t = n - 1$ to $t = n$ days. (**E**) Allografts show lower $(\overline{T}_{5/6} - \overline{T}_{4/5})$ than isografts (n = 5). PAS-stained histological sections show that at t ~ 4 days, (**F**) the isograft kidney is normal whereas (**G**) the allograft kidney shows signs of type I acute rejection. (**H**) BUN and (**I**) serum creatinine at t ~ 4 days are not significantly different (ns) between allografts and isografts (**J**) $(\overline{T}_4 - \overline{T}_3)$ is a statistically significant metric for the T_{kidney} bump in Fig. 2D (n = 5). The green shaded region in Fig. 3, E, F, H, and I represent normal levels for Lewis rats without transplant.

diffuse interstitial inflammation and tubulitis. All five medicated allografts (FK 1 mg/kg) display histological evidence of rejection at the end points (Fig. 4, D to H).

Early indication of acute rejection

The bump observed for data from medicated allografts can be characterized by a statistically significant temperature rise $(\overline{T_{14}} - \overline{T_{10}})$ and decline $(\overline{T_{20}} - \overline{T_{14}})$ (median rise ~0.15°C, P = 0.0353, median decline approximately −0.3°C, P = 0.0184) compared to isografts (Fig. 5A,B), and are "muted" relative to the unmedicated allografts (~0.6°C). Blood markers fail to detect rejection until $t \geq 27$ days. This time point is much later than the onset of rejection, which occurs between t = 10 and 14 days as evidenced by histology (Fig. 4B and 4I).

Fourier transform analysis of T_{kidney} for t = 7 to 21 days (corresponding to the post-medication period, inflection point, and coinciding with the muted bump) reveals the presence of strong ultradian ($f > 1$ day^{-1}) rhythms (specifically, $f = 2$ day^{-1}) for all medicated allografts relative

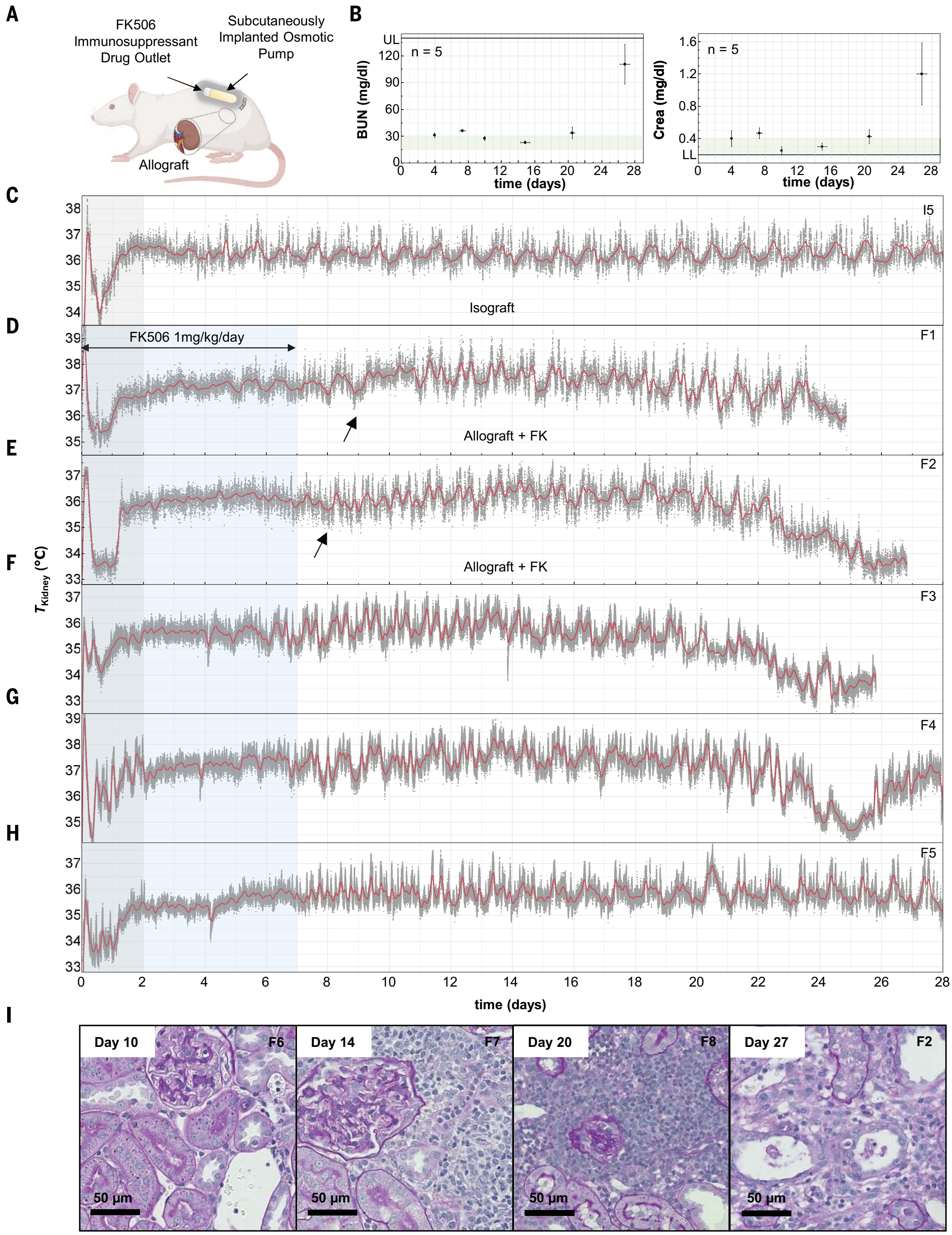


Fig. 4. Delayed rejection after discontinuation of immunosuppressants. (**A**) Illustration of the delivery of FK506 to an allograft at a continuous dose of 1 mg/kg per day for t = 0 to 7 days. (**B**) Serum creatinine and BUN collected at discrete time points (t ~ 4, 7, 10, 14, 20, and 27 days) for animals treated with FK506. (**C**) T_{kidney} versus t for a representative isograft for 28 days. (**D** to **H**) T_{kidney} versus t for n = 5 allografts treated with 1 mg/kg per day FK506 for t = 0 to 7 days (treatment period denoted by blue shaded region). The gray shaded region (t = 0 to 2 days) indicates the post-surgery recovery period. The gray points represent the raw data. The red curve is a smoothing spline fit (λ = 0.01). The black arrows correspond to the onset of a muted temperature bump. (**I**) Representative histological sections stained with PAS for a harvested kidney at t = 10 days shows normal kidney parenchyma without any features of acute rejection whereas those at t = 14, 20, and 27 days all demonstrate acute type I tubulointerstitial rejection, characterized by diffuse interstitial inflammation and frequent tubulitis.

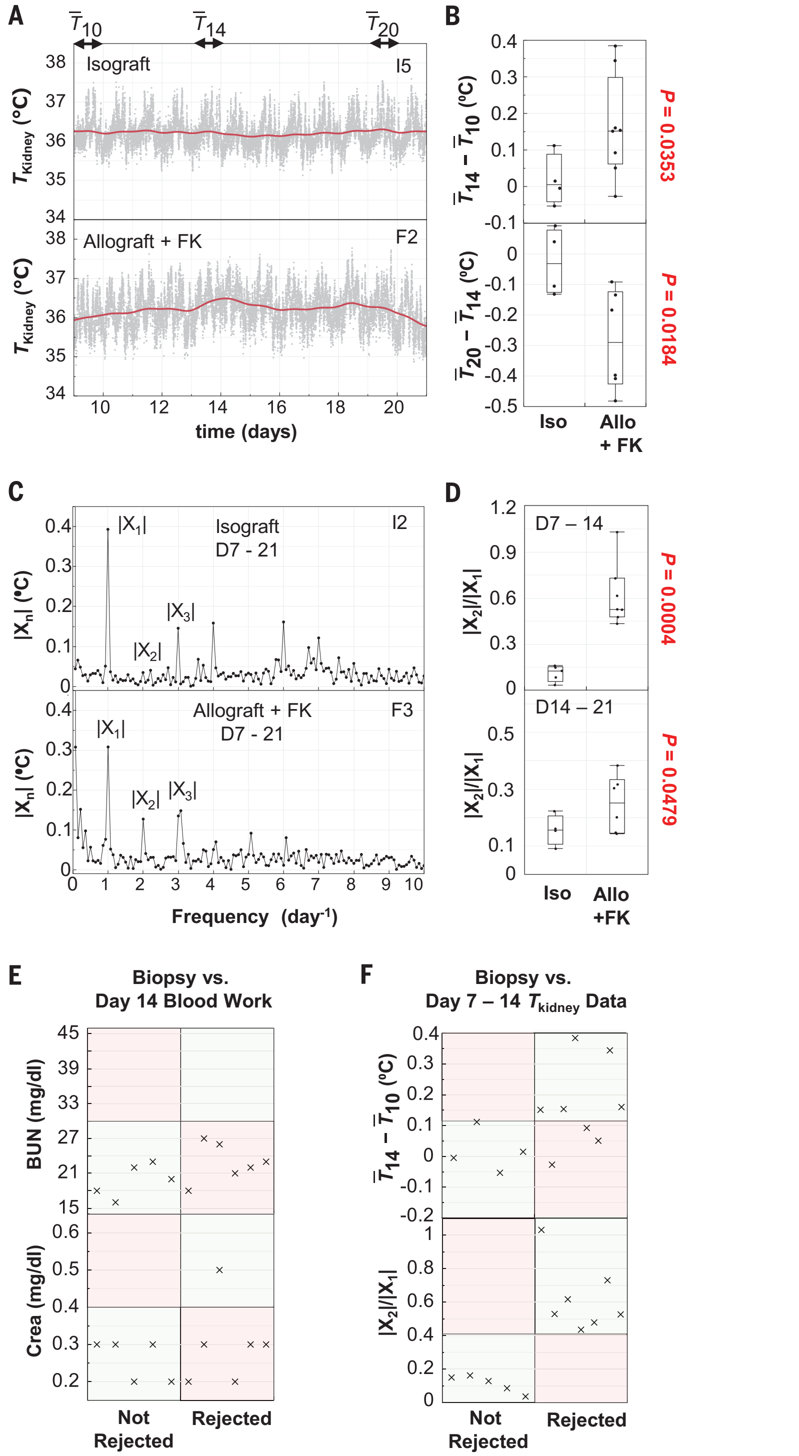


Fig. 5. Temperature as an early indicator of rejection relative to blood markers. (**A**) T_{kidney} for a representative isograft and allograft for t = 9 to 21 days. Gray points represent the raw data. The red curve is a smoothing spline fit (λ = 30). (**B**) The bump rise ($\overline{T}_{14}$ to $\overline{T}_{10}$) and subsequent decline ($\overline{T}_{20}$ to $\overline{T}_{14}$) are significant in medicated allografts compared to isografts (n = 5). (**C**) Fourier Transform of the T_{kidney} data from fig. S14 and Fig. 4F. (**D**) The ratio of the magnitudes of the f = 2 day^{-1} $|X_2|$ and f = 1 day^{-1} ($|X_1|$) peaks for t = 7 to 14 days (top) and t = 14 to 21 days (bottom) is greater in medicated allografts relative to isografts (n = 5). (**E**) Confusion matrices comparing the measured values of BUN and creatinine in isografts and medicated allografts to histological diagnosis. (**F**) Confusion matrix comparing ($\overline{T}_{14}$ to $\overline{T}_{10}$) and $\frac{|X_2|}{|X_1|}$ to blinded histological diagnosis. The cutoff values in (E) are the upper limits of normal BUN and Creatine for Lewis rats and in (F) are the corresponding grand means of the datasets. The red shaded regions correspond to false-positive and -negative evaluations, and the green shaded regions are true negative and positive outcomes.

to isografts (Fig. 5C). The ratio of amplitudes of the half-day and circadian cycles ($\frac{|X_2|}{|X_1|}$) is statistically significant between t = 7 to 14 days (P = 0.0004) and remains so between t = 14 to 21 days (P = 0.0479) (Fig. 5D). Similarly, ($\frac{|X_3|}{|X_1|}$) is also statistically significant (fig. S31). The presence of stronger ultradian features in medicated allografts may be a result of the cyclic nature of T cell activity and/or cellular repair/damage processes, synced with the circadian clock (*28*–*30*).

The subclinical relevance of T_{kidney} and creatinine/BUN can be established by comparison with histology. Accuracy and true positive rate (TPR) calculated from confusion matrices of BUN (accuracy = 46%, TPR = 0%) and creatinine levels (accuracy = 54%, TPR = 17%) illustrate that blood markers cannot detect early-stage or onset of rejection (t = 14 days); nearly all medicated allografts result in false negatives (Fig. 5E). On the other hand, features in T_{kidney} such as the bump (accuracy = 75%, TPR = 62%) and the half-day ultradian rhythm (accuracy = 100%, TPR = 100%) correctly identify early-stage or onset of rejection (Fig. 5F).

Conclusions

This study in rat transplantation models establishes biophysical measurements of T_{kidney} and k_{kidney} as a technique to detect subclinical acute rejection. Compared to invasive biopsies, these biosensors offer dense, real-time, long-term information about surgical recovery, impact of medications, circadian/ultradian rhythms, motion/activity, and graft rejection. T_{kidney} provides early warning of rejection and greater sensitivity relative to serum creatinine and BUN, ~2 to 3 weeks and ~3 days in advance in cases of discontinued and absent administration of immunosuppressive therapy, respectively. Such measurements may help guide personalized dosing strategies and in understanding the efficacy of immunosuppressants.

REFERENCES AND NOTES

1. K. L. Lentine *et al.*, *Am. J. Transplant.* **22** (suppl. 2), 21–136 (2022).
2. L. D. Cornell, R. N. Smith, R. B. Colvin, *Annu. Rev. Pathol.* **3**, 189–220 (2008).
3. U.S. Department of Health and Human Services, National Institutes of Health, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, "2021 Annual Data Report" (USRDS, 2021); https://usrds-adr.niddk.nih.gov/2021.
4. T. Y. Kee *et al.*, *Transplantation* **82**, 36–42 (2006).
5. D. N. Rush, S. F. Henry, J. R. Jeffery, T. J. Schroeder, J. Gough, *Transplantation* **57**, 208–210 (1994).
6. W. Zhang *et al.*, *J. Am. Soc. Nephrol.* **30**, 1481–1494 (2019).
7. R. Mehta *et al.*, *Am. J. Transplant.* **18**, 1710–1717 (2018).
8. K. Bakdash *et al.*, *Semin. Intervent. Radiol.* **36**, 97–103 (2019).
9. D. J. Lo, B. Kaplan, A. D. Kirk, *Nat. Rev. Nephrol.* **10**, 215–225 (2014).
10. C. L. Edelstein, in *Biomarkers of Kidney Disease (Second Edition)*, C. L. Edelstein, Ed. (Academic Press, 2017), pp. 241-315.
11. Q. D. Mac *et al.*, *Nat. Biomed. Eng.* **3**, 281–291 (2019).
12. P. J. O'connell, K. L. Keung, M. C. Menon, B. Murphy, in *Kidney Transplantation - Principles and Practice (Eighth Edition)*, S. J. Knechtle, L. P. Marson, P. J. Morris, Eds. (Elsevier, 2019), pp. 418-433.
13. R. D. Bloom *et al.*, *J. Am. Soc. Nephrol.* **28**, 2221–2232 (2017).

14. J. J. Friedewald *et al*., *Am. J. Transplant.* **19**, 98–109 (2019).
15. R. El Fekih *et al*., *J. Am. Soc. Nephrol.* **32**, 994–1004 (2021).
16. M. Naesens, D. Anglicheau, *J. Am. Soc. Nephrol.* **29**, 24–34 (2018).
17. E. Van Loon *et al*., *Am. J. Transplant.* **21**, 740–750 (2021).
18. M. Suthanthiran *et al*., *N. Engl. J. Med.* **369**, 20–31 (2013).
19. J.-L. Gennisson *et al*., "Multiwave technology introducing shear wave elastography of the kidney: Preclinical study on a kidney fibrosis model and clinical feasibility study on 49 human renal transplants" in *2010 IEEE International Ultrasonics Symposium*, pp. 1356-1359 (2010).
20. C. A. Romero *et al*., *Am. J. Physiol. Renal Physiol.* **314**, F99–F106 (2018).
21. R. N. Raman *et al*., *J. Biomed. Opt.* **14**, 020505 (2009).
22. J. Schneiderman *et al*., *Sci. Rep.* **12**, 7298 (2022).
23. J. C. Doloff *et al*., *Nat. Biomed. Eng.* **5**, 1115–1130 (2021).
24. O. Laitano *et al*., *Shock* **50**, 226–232 (2018).
25. S. Ebong *et al*., *Infect. Immun.* **67**, 6603–6610 (1999).
26. K. D. Vlach, J. W. Boles, B. G. Stiles, *Comp. Med.* **50**, 160–166 (2000).
27. L. L. Monroe *et al*., *Shock* **46**, 723–730 (2016).
28. M. Perelis, M. L. Ford, X. Luo, *Am. J. Transplant.* **15**, 293 (2015).
29. X. Yu *et al*., *Science* **342**, 727–730 (2013).
30. G. Fernandes, in *Biologic Rhythms in Clinical and Laboratory Medicine,* Y. Touitou, E. Haus, Eds. (Springer, 1992), pp. 493-503.

ACKNOWLEDGMENTS

This work made use of the NUFAB facility of Northwestern University's NUANCE Center, which has received support from the SHyNE Resource (NSF ECCS-2025633), the IIN, and Northwestern's MRSEC program (NSF DMR-1720139). This work made use of the Microsurgery and Preclinical Research Core Facility in the Comprehensive Transplant Center at the Feinberg School of Medicine. The authors thank J. Zhu, S. Coughlin, and J. Winograd for preliminary efforts in microfabrication. The authors thank A. Spann (Center for Advanced Molecular Imaging, Northwestern University) for 3D reconstruction of and segmentation of the contrast-CT image. The authors thank D. Procissi and N. Bertolino (Small Animal Imaging Facility, Center for Translational Imaging, Northwestern University) for collection of CT images. We thank M. Arceta, J. Hernandez, and S. Popovics for assistance with animal husbandry and care (Center for Comprehensive Medicine, Northwestern University). The authors acknowledge the Northwestern Mouse Histology and Phenotyping Laboratory for preparation of histological samples and the Northwestern Pathology Core Facility for assistance with scanning of histology slides. Figures 2A, 2B, and 4A were created using Biorender.com. Illustrations in Fig. 1A were created by J. Sinn-Hanlon, iLearning@VetMed, UIUC. The authors thank J. Ciatti for the photograph in Fig. 1C and for performing surface roughness measurements. **Funding:** S.R.M. acknowledges support from the NSF Graduate Research Fellowship (NSF DGE-1842165). M.P. acknowledges support in part by an Alpha Omega Alpha Carolyn L. Kuckein Student Research Fellowship. **Author contributions:** Project conceptualization: S.R.M., J.A.R., and L.G. Design of hardware/software: S.R.M. Thermal sensor fabrication/calibration: S.R.M. Theoretical simulations: H.W. with guidance from Y.H. Design of Experiments: S.R.M., J.A.R., Z.J.Z., and L.G. Data analysis: S.R.M. with review by J.A.R., Z.J.Z., and L.G. Surgeries: J.-J.W. and X.Z. with assistance from S.R.M. Animal blood draws/small procedures: J.-J.W. with assistance from S.R.M. and M.P. Post-operative care, monitoring of animals: S.R.M., M.P., and J.-J.W. Blinded Histology Review and Diagnosis: A. C. Project Supervision: J.A.R. and L.G. Original manuscript draft and figure preparation: S.R.M. Draft editing: S.R.M., J.A.R., L.G., and Z.J.Z. All authors read and provided comments on the manuscript. **Competing interests:** J.A.R., S.R.M., L.G., and Z.J.Z. have filed for a patent based on the research in this manuscript (PCT/US2023/010402). **Data availability:** All data are available in the manuscript or the supplementary materials.

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh7726
Materials and Methods
Figs. S1 to S32
Tables S1 and S2
References (*31–34*)
MDAR Reproducibility Checklist

Submitted 13 March 2023; accepted 19 July 2023
10.1126/science.adh7726

NEUROSCIENCE

Lifelong restructuring of 3D genome architecture in cerebellar granule cells

Longzhi Tan[1,2]*†, Jenny Shi[1,2,3]†, Siavash Moghadami[1,4], Bibudha Parasar[1], Cydney P. Wright[1,5], Yunji Seo[1], Kristen Vallejo[6], Inma Cobos[6], Laramie Duncan[7], Ritchie Chen[2], Karl Deisseroth[2,7,8]*

The cerebellum contains most of the neurons in the human brain and exhibits distinctive modes of development and aging. In this work, by developing our single-cell three-dimensional (3D) genome assay—diploid chromosome conformation capture, or Dip-C—into population-scale (Pop-C) and virus-enriched (vDip-C) modes, we resolved the first 3D genome structures of single cerebellar cells, created life-spanning 3D genome atlases for both humans and mice, and jointly measured transcriptome and chromatin accessibility during development. We found that although the transcriptome and chromatin accessibility of cerebellar granule neurons mature in early postnatal life, 3D genome architecture gradually remodels throughout life, establishing ultra–long-range intrachromosomal contacts and specific interchromosomal contacts that are rarely seen in neurons. These results reveal unexpected evolutionarily conserved molecular processes that underlie distinctive features of neural development and aging across the mammalian life span.

Different cell types within the same organism can mature along highly distinctive developmental and aging trajectories. At the molecular level, cell type–specific gene transcription can be orchestrated by diversity in genome architecture (folding of chromosomes in three dimensions) (*1*). Yet the genome architecture over the life span has not been elucidated, limiting our understanding of the life-spanning cellular dynamics of brain function and dysfunction.

Previously, by using our single-cell three-dimensional (3D) genome assay (diploid chromosome conformation capture, or Dip-C) (*2*–*4*), we found that cells in the mouse forebrain (cerebral cortex and hippocampus) undergo cell type–specific transformation in transcriptome and genome architecture during the first month of life (*5*). However, technological limitation hindered generalization of this finding across brain regions, species, and life span. In this work, we broadened our perspective along all three of these dimensions by turning our focus across the neuraxis—from forebrain to hindbrain—specifically to the cerebellum, which contains ~80% of all neurons in the human brain. The cerebellum, a powerful yet compact processing unit that has expanded over evolution (*6*), exhibits distinct characteristics, including prolonged development after birth, malformation in autism (*7*), and degeneration during aging (*8*). Understanding genome dynamics of the cerebellum may provide insights into these features, as well as into motor control and cognition (*9*).

Prior work has revealed nuclear morphological features of cerebellar cells in vitro (*10*), but the 3D genome structures remain to be fully solved. Recently, chromosome conformation capture (3C or Hi-C) was performed on the adult cerebellum (*11*–*13*); however, a comprehensive, cross-species, single-cell 3D genome atlas of the developing and aging cerebellum is lacking. In addition, simultaneous analysis of transcriptome (*14*) and chromatin accessibility (*15*) in the cerebellum would provide additional valuable information. In this work we show that the cerebellum undergoes an extraordinary, lifelong 3D genome transformation that is conserved between human and mouse and is far greater in magnitude than that of the forebrain (*5*), revealing genome rewiring as a potential molecular hallmark of aging.

A 3D genome atlas of the developing and aging cerebellum

Granule cells (the vast majority of cerebellar neurons) are generated between postnatal days 0 and 21 (P0 and P21) in mice and between the third trimester of pregnancy and ~1 year after birth in humans, which is much later than the same process in the cerebral cortex (*5*). During this period, granule cell progenitors divide and migrate from the external granular layer (EGL) to the internal granular layer (IGL), expanding more than 100-fold in number. Toward the other end of the life span, the cerebellum is also known to exhibit a slow epigenetic aging clock of DNA methylation (*16*). To explore the genomic underpinnings of this entire timeline, we created a 3D genome atlas that extends across the human and mouse life span, alongside a multi-ome atlas focused on human development (Fig. 1A).

We first simultaneously profiled transcriptome and chromatin accessibility during postnatal development of the human cerebellum, sequencing 63,768 cells from seven donors (*17*) (one adult and six between the ages of 0.1 and 2.3 years) and detecting a median of 645 to 1617 genes [944 to 3845 unique molecular identifiers (UMIs)] and 12,000 to 34,000 ATAC (assay for transposase-accessible chromatin) fragments per cell from each donor (tables S1 and S2). We additionally profiled a critical age in mouse—P14, when cells are present in both the EGL and the IGL—sequencing 7182 cells and detecting a median of 618 genes (944 UMIs) and 22,000 ATAC fragments per cell.

We then comprehensively profiled 3D genome architecture across the human and mouse life span, sequencing 11,207 cells (Fig. 1A). For humans, we sequenced 5202 cells from 24 donors (0.1 to 86 years) and obtained a median of 608,000 chromatin contacts per cell. Among these cells, 3580 came from the cerebellum (chiefly lateral; vermis if lateral was not available), and 1622 from the cerebral cortex [Brodmann area (BA) 46 of the dorsolateral prefrontal cortex (DLPFC)] of the same donors (tables S2 to S4). For mice, we sequenced 6005 cells from the cerebellum (birth to 21 months), obtaining a median of 496,000 contacts per cell, and incorporated our prior dataset of 1075 and 879 cells from the cerebral cortex and hippocampus, respectively (*5*).

Transcriptionally immature granule cells in the newborn human cerebellum

We integrated the transcriptome dataset from all seven human donors using the LIGER (linked inference of genomic experimental relationships) software (Fig. 1B) (*14*, *18*, *19*); modifying LIGER parameters did not affect conclusions (fig. S1). Granule cells were the predominant cell type at all ages examined (median, 83%; range, 76 to 92%). Astrocytes (median, 7%; range, 3 to 11%) were composed of the cerebellum-specific Bergmann glia (4%) (*18*) and the typical parenchymal astrocytes (2%) (*14*, *15*) (Fig. 1C). We further identified gene *GABRG3* as a specific marker for Bergmann glia (Fig. 1C). Other cell types included molecular layer interneurons (MLIs) (4%), oligodendrocytes (3%), microglia (1%), and unipolar brush cells (UBCs) (0.3%). Purkinje cells were too rare to be reliably identified.

Unlike the case of mouse brain—in which granule cells are almost entirely in the EGL at birth—the human cerebellum is already dominated at birth by IGL neurons. However, it remains unknown when and how human granule cells mature at the genomic level. We found that at birth, granule cells were subdivided into maturation stages by transcriptomic measures (Fig. 1B). In our integrated transcriptome data,

[1]Department of Neurobiology, Stanford University, Stanford, CA 94305, USA. [2]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. [3]Department of Chemistry, Stanford University, Stanford, CA 94305, USA. [4]Department of Chemical and Systems Biology, Stanford University, Stanford, CA 94305, USA. [5]Department of Biology, Stanford University, Stanford, CA 94305, USA. [6]Department of Pathology, Stanford University, Stanford, CA 94305, USA. [7]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA 94305, USA. [8]Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305, USA.
***Corresponding author. Email: deissero@stanford.edu (K.D.); tttt@stanford.edu (L.T.)**
†These authors contributed equally to this work.

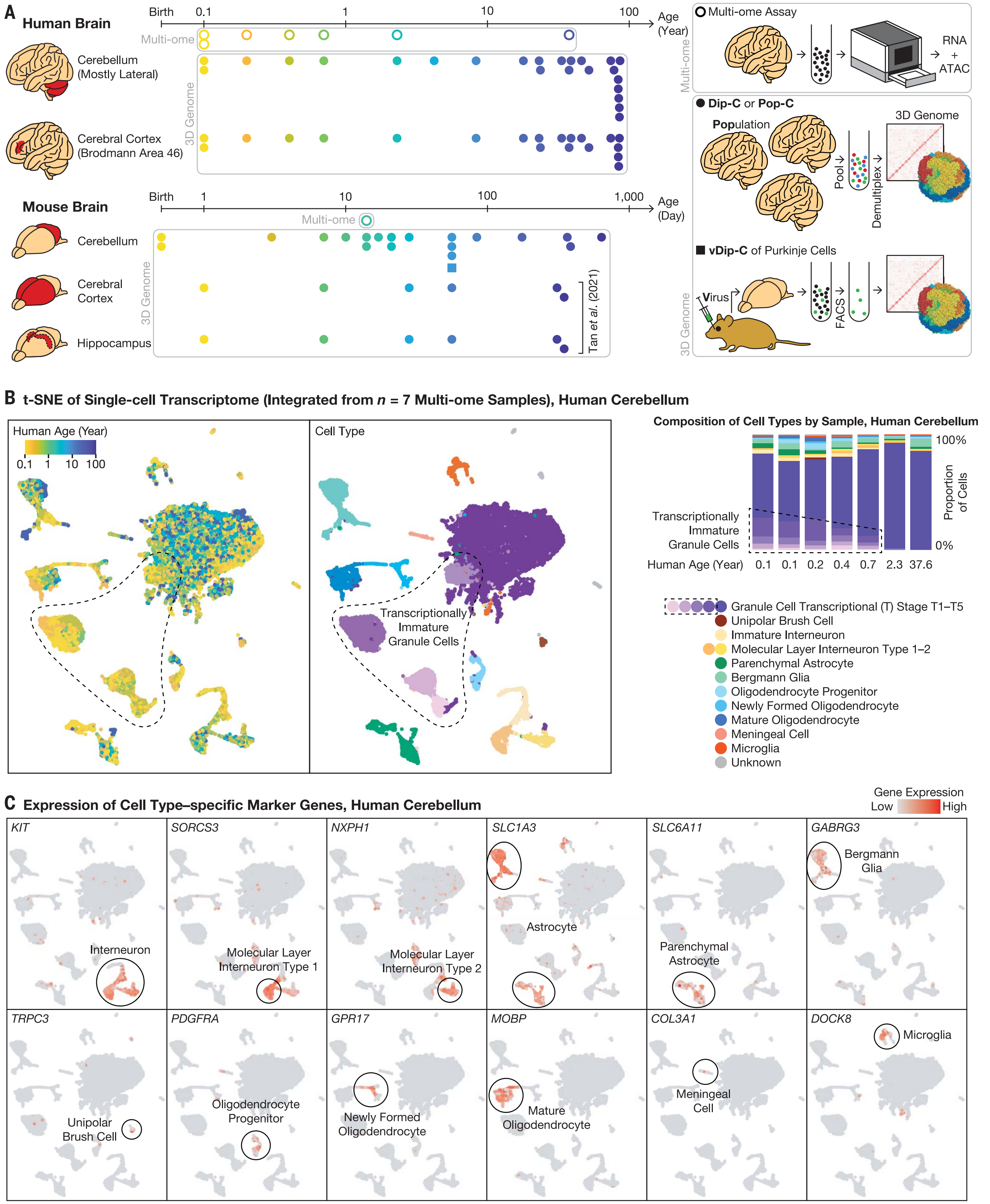


Fig. 1. 3D genome atlas across life span for human and mouse cerebellum with multi-ome atlas of postnatal development. (**A**) Study design. (**B**) Integrative transcriptome analysis of human multi-ome samples. t-SNE, t-distributed stochastic neighbor embedding. (**C**) Representative expression profiles of marker genes.

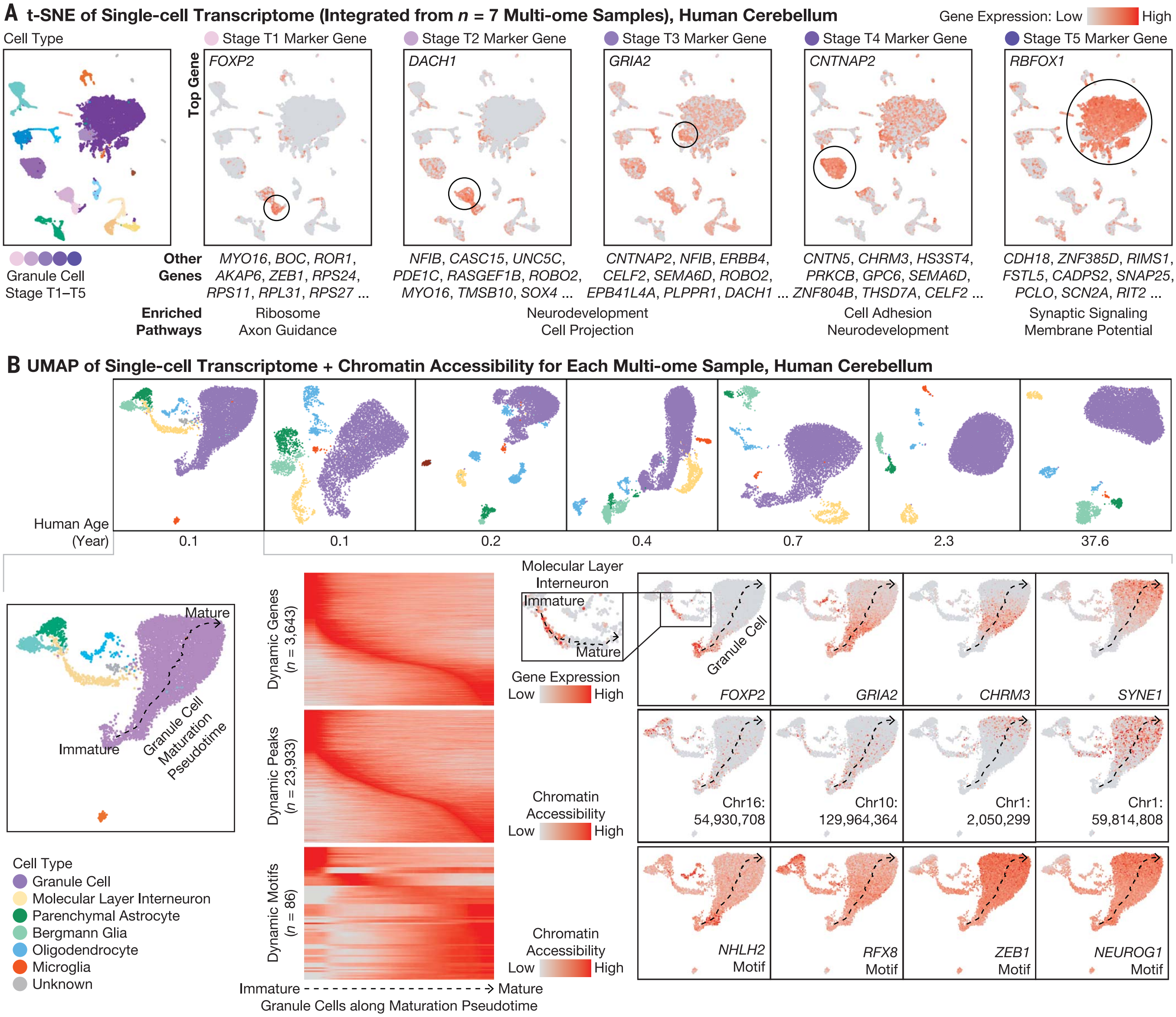


Fig. 2. Simultaneous transcriptome and chromatin accessibility profiling revealed continuous maturation of cerebellar granule cells over the first postnatal year. (**A**) Stages of granule cell maturation, their marker genes (ranked by specificity), and enriched pathways (summarized for the top 100 genes). (**B**) Maturation pseudotime analysis of each sample. *n* = number of dynamic genes, peaks, and motifs. Reference genome, hg38.

granule cells existed in one mature form (termed transcriptional stage T5) and several immature forms (T1 to T4). T5 was the predominant form (99 to 100%) in the 2.3- and 37.6-year-old donors. In younger donors, however, T1 to T4 made up a substantial fraction: 32 to 34% in the 0.1 and 0.2 year olds, 23% in the 0.4 year old, and 14% in the 0.7 year old (Fig. 1B), revealing an abundance of transcriptionally immature granule cells.

T1 to T5 granule cells expressed partially overlapping sets of genes (Fig. 2A and table S5). T1 was enriched for ribosomal subunits [false discovery rate (FDR) < 10^{-10}; including *RPS24/11/27* and *RPL31/39/32*] as well as axon guidance–related genes (FDR = 1×10^{-4}; including *BOC* and *LAMA2*) and expressed *FOXP2*. T2 was enriched for neurodevelopmental genes (FDR = 3×10^{-6}; including *NFIB* and *UNC5C*) including additional genes related to morphogenesis of projections (FDR = 6×10^{-10}; including *ROBO2* and *MYO16*). T3 was also enriched for neurodevelopment (FDR < 10^{-10}; including *ERBB4*) and projection morphogenesis (FDR = 1×10^{-9}; including *SEMA6D*) and expressed *GRIA2*. T4 was enriched for cell adhesion (FDR < 10^{-10}; including *CNTNAP2/5* and *CNTN5*) and genes related to neurodevelopment (FDR = 2×10^{-9}; including *CHRM3* and *GPC6*). T5 was enriched for synaptic signaling genes (FDR < 10^{-10}; including *CADPS2* and *SNAP25*) and regulation of membrane potential (FDR < 10^{-10}; including *RIMS1* and *SCN2A*) and expressed *RBFOX1*. Correlated gene module analysis confirmed these genes and pathways (fig. S3 and table S5).

A continuum of granule cells and interneurons with maturing transcriptome and chromatin accessibility

For each donor, we jointly analyzed transcriptome and chromatin accessibility using ArchR (*20*) (Fig. 2B). Despite discrete appearances in LIGER (Fig. 1A), granule cells formed a continuous developmental pseudotime for each donor below the age of 1 year (figs. S4 and S6A). Dynamically expressed genes included many genes from LIGER analysis. Dynamically accessible transcription factor–binding motifs included ASCL1/2 and NHLH1/2 (early); KLF11/14 and RFX2/3/4/8 (intermediate); and

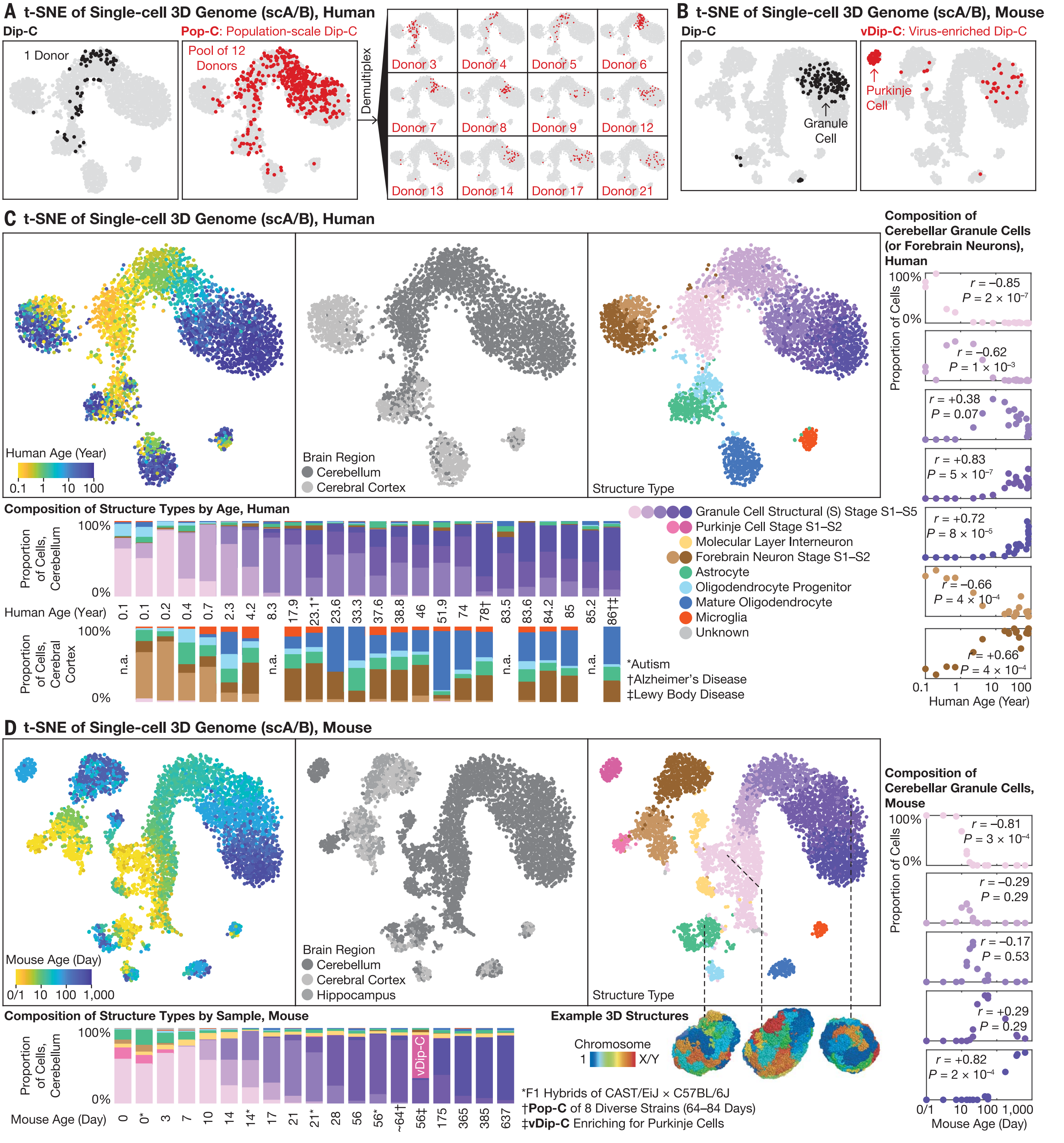


Fig. 3. High-throughput, high-precision 3D genome profiling uncovered lifelong genome remodeling in the human and mouse cerebellum. (**A**) Pop-C method. (**B**) vDip-C method. (**C** and **D**) Cross-species 3D genome atlas for the developing and aging cerebellum (with cerebral cortex as counterpoint). Pearson's *r* (and *P* value) was calculated from logarithm of age.

NEUROG1/2/3, NEUROD4/6, ZEB1, MEF2A/B/C/D, and NFIA/B/X (late). We thus found the newborn human cerebellum to include a complex mixture of granule cells with continuously evolving transcriptomic and epigenomic states. We validated aspects of this continuum by reanalyzing published transcriptome-only data (fig. S6B) (*14*). A continuum was also observed in mouse cells (fig. S7) (*14*).

We observed a similar continuum of maturation in MLIs (fig. S8). Immature MLIs were abundant at birth and vanished over age (64 to 71% in the 0.1 and 0.2 year olds, 44% in the 0.4 year old, 25% in the 0.7 year old, 5% in the 2.3 year old, and <1% in the adult) (Fig. 1B). In addition, immature MLIs and immature granule cells shared expression of many genes, such as *FOXP2* (Fig. 2B).

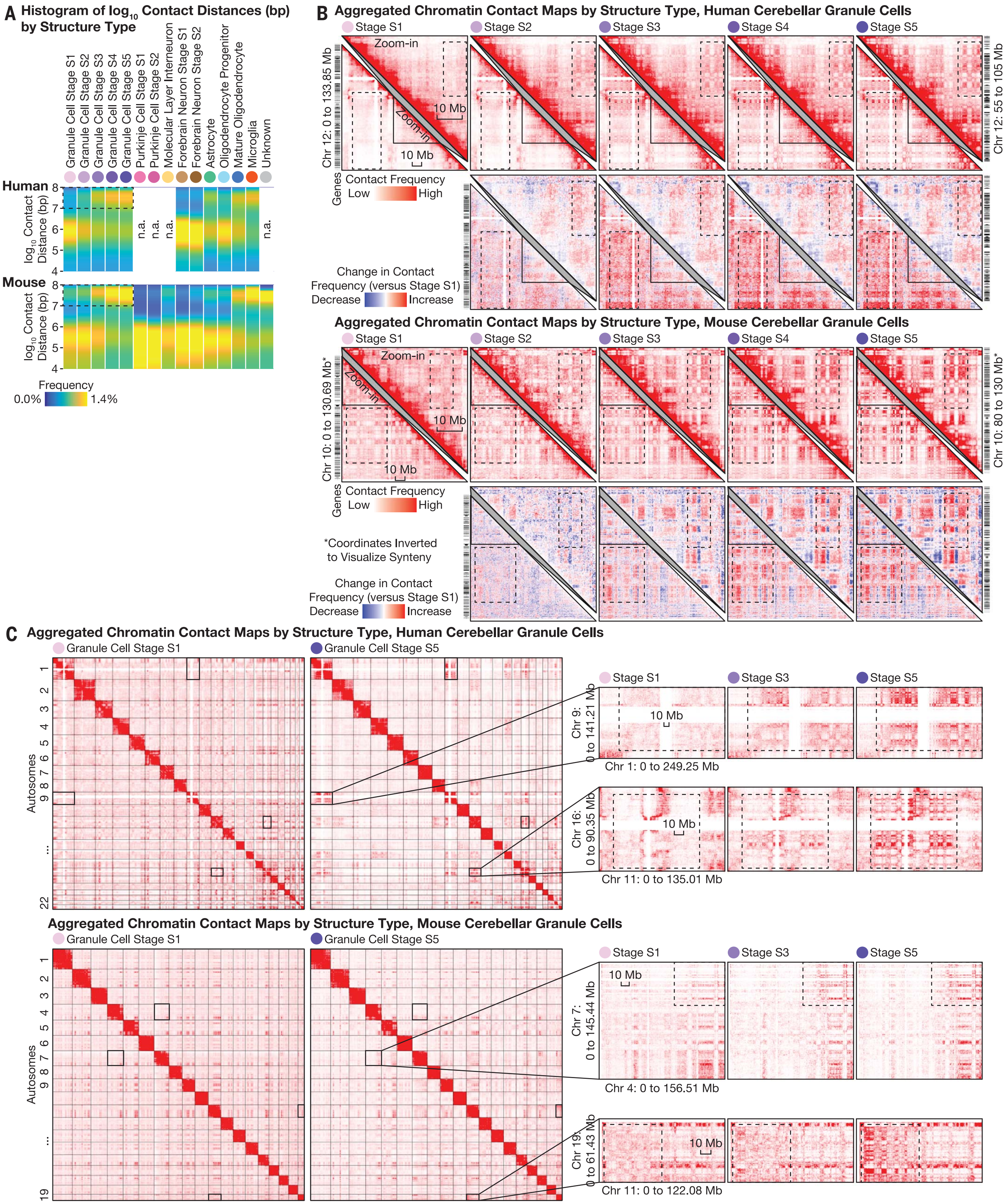


Fig. 4. Cerebellar granule cells formed ultra–long-range intrachromosomal contacts and specific interchromosomal contacts during development and aging. (**A**) Distribution of genomic distances of chromatin contacts. (**B**) Aggregated contact maps for an example chromosome and a zoomed-in region (upper-right triangles). Zoomed-in regions are homologous. Dashed boxes highlight prominent changes. Bin size, 250 kb. Reference genomes, hg19 and mm10. (**C**) Aggregated interchromosomal contact maps. Bin sizes are 6 Mb (human), 5 Mb (mouse), and 500 kb (zoomed-in regions).

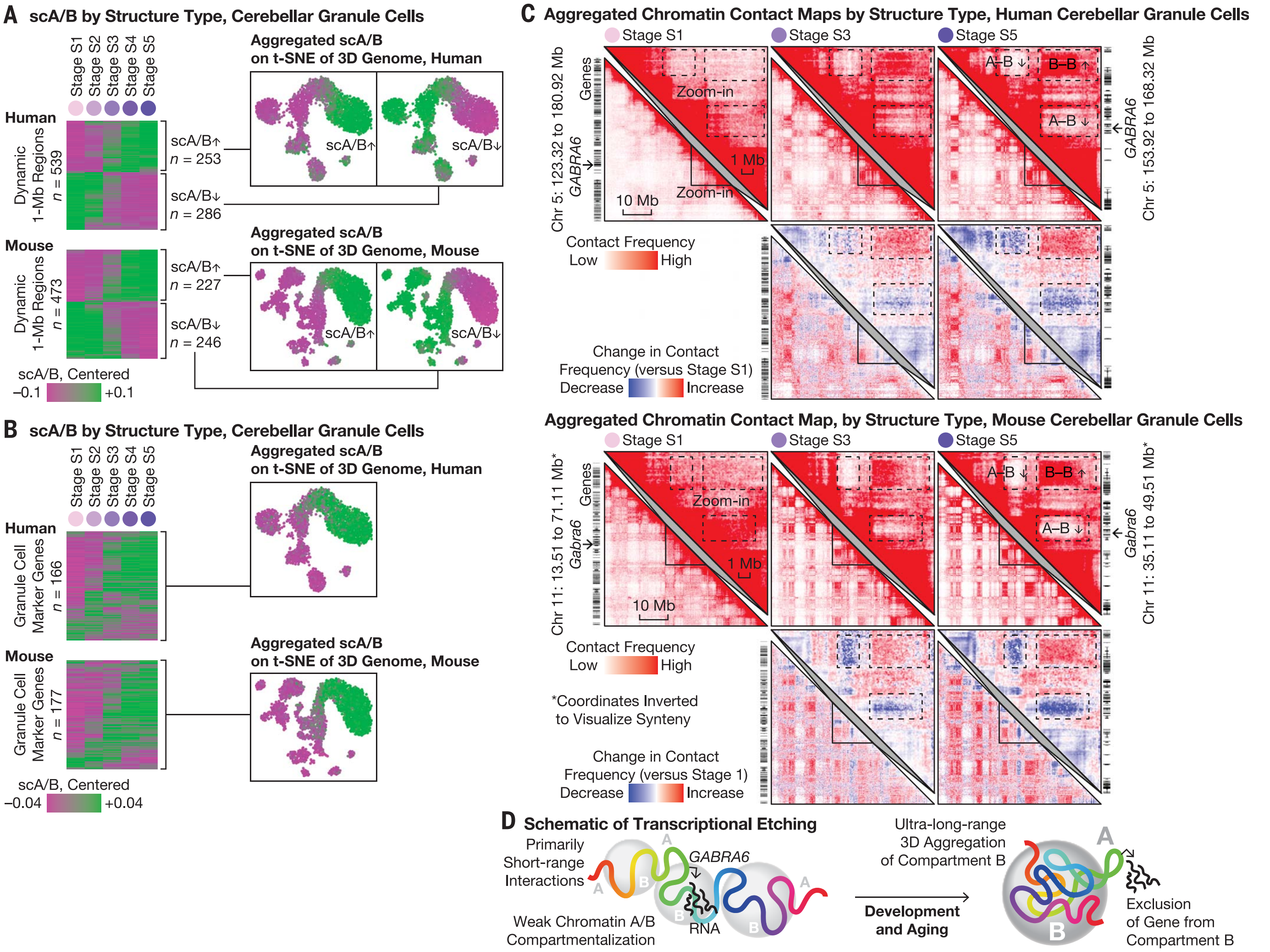


Fig. 5. Lifelong maturation of chromatin A/B compartments was associated with cerebellar granule cell–specific genes. (**A**) Mean scA/B of each dynamic 1-Mb region at each stage (left). Rows are ordered by hierarchical clustering; the two clusters were visualized on t-SNE plots (right). As in (*5*), scA/B calculation excludes contacts within each region and thus primarily reports on long-range interactions. *n* = number of 1-Mb regions. (**B**) (Left) Mean scA/B of each 1-Mb region harboring granule cell–specific marker genes (*14*) at each stage, (right) with aggregated scA/B shown on t-SNE plots. (**C**) Aggregated contact maps for an example gene. Bin size is 100 kb. Reference genomes, hg19 and mm10. (**D**) Schematic of transcriptional etching.

3D genome profiling of diverse populations and rare cells with Pop-C and vDip-C

We next focused on genome architecture. Cerebellar cells exhibit distinct genome morphology, beginning with nuclear dimensions: Granule cell nuclei are among the smallest in the brain (5 to 6 μm in diameter), whereas Purkinje cells have large nuclei (~12 μm in diameter) (*21*). During differentiation, cultured mouse granule cell progenitors reduce nuclear volume and spatially redistribute histone H3.3 (*10*). However, 3D genome structures of cerebellar cells remain unclear, and little is known about lifetime-spanning dynamics in vivo.

To meet this challenge, we developed two 3D genome technologies. First, population-scale Dip-C (Pop-C) leverages the whole-genome sequencing capability of Dip-C (*2*) to pool a large number of samples and computationally demultiplexes (*22*) cells based on natural genetic variations, with high genomic coverage of 10 to 20% (Figs. 1A and 3A). We validated Pop-C by computationally pooling known samples (accuracy, 672/672 = 100%) (fig. S10).

In the simplest case, many of our mouse samples were a pool of males and females, which we demultiplexed based on the ratio of reads between the X chromosome and autosomes (fig. S9). In a more complex case, we pooled one mouse each from the eight founder strains of the JAX Diversity Outbred (DO) collection (*23*) and demultiplexed based on known single-nucleotide polymorphisms (SNPs) (table S4). In the most challenging case, we pooled 3 to 13 unrelated human individuals and demultiplexed them (*22*) based on common SNPs among populations without prior knowledge about donor genotypes (Fig. 3A). Pop-C was thereby shown to provide a robust method for profiling single-cell 3D genome structures at scale (fig. S11).

The second method we developed, vDip-C, enables genomic profiling of rare cell populations without the use of transgenic mouse lines (Figs. 1A and 3B). We used a single viral vector containing a cell type–specific promoter; an ultrabright, fixation-resistant, monomeric fluorescent protein (*24*); and a nuclear membrane localization sequence (*25*) (fig. S12) and administered it to wild-type mice (*26*) through retro-orbital injection of adeno-associated virus (*27*, *28*).

We used vDip-C to solve the 3D genome structures of mouse Purkinje cells (Fig. 3B). Although Purkinje cells are abundant at birth (P0), they quickly become outnumbered by granule cells (Fig. 3D). To isolate this rare (<0.5%)

cell type from adults, we constructed a vDip-C vector with a Purkinje cell–specific promoter (*Pcp2*) (*29*), administered the viral vector to wild-type mice, and isolated nuclei by fluorescence-activated cell sorting (FACS) (fig. S12).

Lifelong 3D genome transformation of granule cells

We created a high-resolution cross-species single-cell 3D genome atlas and resolved 3D genome structures for a subset of cells (from F_1 hybrid mice) (Figs. 1A, 3C, and 3D). Similar to our previous studies (*2*, *3*, *5*), single-cell chromatin A/B compartment (scA/B) analysis revealed 3D genome structure types that corresponded to diverse cerebellar cell types, including granule cells, astrocytes, oligodendrocytes, and microglia in both species, as well as MLIs and Purkinje cells in mice. Replicates yielded reproducible scA/B and contact patterns (Fig. 3A and figs. S11 and S21). Of our 24 donors, three were diagnosed with autism, Alzheimer's disease, and/or Lewy body disease; excluding them did not affect our conclusions (fig. S28).

Granule cells exhibited by far the most dramatic structural transformation. Granule cells of both species were born with an immature structure type, termed structural stage S1, that resembled forebrain neurons (Fig. 3, C and D, and figs. S14 and S19). As the cerebellum developed and aged, granule cells continuously and progressively evolved into new structure types S2 to S5, which increasingly differed from forebrain neurons (figs. S14 and S19). This transformation was the primary source of scA/B variations (the first principal component) and could be visualized regardless of the analysis method (figs. S15 and S16).

In human cells, abundances of S1 to S5 structures peaked around the ages of 0.2, 1, 10, 30, and 80 years, respectively, although considerable between-donor variability was observed (Fig. 3C). This age distribution suggested that S2 to S5 likely all corresponded to T5. In mouse cells, S1 to S5 peaked around P3, P14, P21, P56, and P365, respectively (Fig. 3D); within S5, the 3D genome continued to mature between P365 or P385 (~12 months) and P637 (~21 months) (fig. S17). These data reveal a 3D genome aging clock of large architectural transformation in a mostly postmitotic cell type.

Ultra–long-range intrachromosomal contacts and specific interchromosomal contacts in granule cells

The most prominent architectural changes in granule cells were the emergence of ultra–long-range (10 to 100 Mb) contacts, which had been thought largely restricted to non-neuronal cells (*13*, *30*), except for mouse rod photoreceptors (Fig. 4A) (*3*). From stage S1 to S5, the fraction of contacts that were ≥10 Mb steadily increased from 19 ± 4 to 33 ± 3% in human cells (mean ± SD; two-sided U test, $P < 10^{-10}$), and from 19 ± 6 to 34 ± 2% in mouse cells ($P < 10^{-10}$), which were far greater in magnitude than for forebrain neurons [from 15 ± 4 to 16 ± 5% during human development (P = 0.038), from 11 ± 4 to 13 ± 3% in mouse cells ($P < 10^{-10}$), and from 9 ± 1 to 10 ± 2% in Purkinje cells ($P = 2 \times 10^{-5}$)] (fig. S18).

This progression in granule cells might be partly driven by their small nuclear size. The abundance of ultra–long-range contacts resembled that seen in non-neuronal cells such as microglia (34 ± 3% in both species) and mature oligodendrocytes (29 ± 4% in human, 27 ± 5% in mouse), both of which have similarly small nuclei. However, these cell types differed substantially in the genomic loci that formed such contacts (fig. S19) and in scA/B profiles (fig. S14), which suggests that nuclear size was not the only driving force.

Granule cells' redistribution of intrachromosomal contacts was accompanied by highly specific interchromosomal contacts. We found increasing interactions among certain human chromosomes, most prominently within a hub of chromosomes (chr) 1, 9, 11, 14, 15, 16, 17, 21, and 22, and between chromosome pairs such as chr 2 and 9, 4 and 14, 8 and 11, and 13 and 20; meanwhile, chr 12 weakened its interactions with the hub (Fig. 4C and fig. S20). Both ultra–long-range intrachromosomal contacts (Fig. 4B) and interchromosomal contacts often involved large stretches of the heterochromatic compartment B interleaved with small stretches of the euchromatic compartment A [also known as "megaloops" or "megaenhancers" (*31*)]. We also observed interchromosomal contacts in mouse cells (Fig. 4C); for example, mouse chr 7 gained interactions with chr 4, 5, 11, 17, and 19. These results highlight another example of conserved, specific interchromosomal interactions beyond prior discoveries in nasal tissue (*3*, *32*–*34*).

Life-spanning scA/B changes associated with granule cell–specific marker genes

Through previous work, we have shown that scA/B generally correlates with cell type–specific gene expression, although discordance can be observed at the single-gene level and regarding temporal dynamics (*3*, *5*). Additionally, it has remained unclear how scA/B interacts with gene expression during aging. In granule cells, we found the predominant mode of scA/B changes to be progressive up- or down-regulation. We calculated the mean scA/B of each 1-Mb genomic region at S1 to S5 and identified the top 20% dynamic regions. In both species, these ~500 dynamic regions either gradually increased or decreased in scA/B across S1 to S5 (Fig. 5A).

We examined genomic regions that harbored conserved marker genes of mature granule cells (*14*). Expression of these ~200 genes began around birth, when T5 emerged. By contrast, on average, these loci gradually increased scA/B throughout life (Fig. 5B and fig. S22). For example, *GABRA6* gradually lost contacts with two nearby heterochromatic regions (which steadily gained contacts with each other) over the life span (Fig. 5C) and consequently increased scA/B until S3 (~10 years) in human cells (two-sided U test, $P < 10^{-10}$ from S1 to S2 and 1×10^{-5} from S2 to S3) and until S4 (~P56) in mouse cells ($P < 10^{-10}$ from S1 to S2, 6×10^{-10} from S2 to S3, and 9×10^{-6} from S3 to S4), persisting well after transcriptional up-regulation at ~0.5 years and ~P10, respectively (*14*). Down-regulated genes on average exhibited relatively unchanged scA/B, which was consistent with previous observations (*3*). Increased transcription may continue to etch 3D genome structure long after initial gene activation (Fig. 5D) (*5*).

Robust 3D genome maturation despite functional perturbations

To test the robustness of this genome restructuring, we explored functional perturbation of chromatin remodeling (fig. S26). Using bulk Dip-C, we observed little effect on 3D genome maturation in mice with clinically relevant heterozygous deletion of *Arid1b* (*35*) or *Chd8*, although we cannot rule out more-subtle differences. Granule cell–specific, homozygous deletion of *Chd4* caused moderate 3D changes (*12*); however, these changes had little overlap with (and were much smaller than) our observed architectural maturation.

Discussion

Once born, most neurons must last for a lifetime; however, we know little about how underlying genomic information may be structurally organized. In this work, we discovered distinctive genome architecture in cerebellar granule cells: ultra–long-range contacts that are uncommon in neurons, specific interchromosomal contacts reminiscent of those in nasal tissue (*3*), and restructuring over decades that may be stabilized by cell type–specific gene transcription. We showed that the mouse is an excellent animal model of this process, despite substantial differences from humans in life span.

We provided mechanistic insights into the principles of this reorganization. For example, both granule cells and Purkinje cells lack neuron-specific non-CpG DNA methylation (*13*), revealing that non-CpG methylation was neither required for our previously discovered neuron-specific radial genome movement (*5*)—which we observed in both cell types (fig. S27) (*36*)—nor required for suppressing ultra–long-range contacts.

A potential function of this architecture might be to manage space and energy expenditure. Human brains are 80% cerebellar granule cells by neuron number. If each granule cell consumed the same volume and energy as a typical neuron in the cerebral cortex, metabolic costs could become prohibitive. Consistent with this idea, granule cells are quiet by firing rate (~0.1 Hz) (*37*) compared with

Purkinje cells (~50 Hz) (*38*). Granule cells might therefore have adopted an energy-saving state physiologically, transcriptionally, and architecturally. Cerebellar and hippocampal granule cells adopt different structural strategies despite having similar nomenclature. Hippocampal granule cells were more similar to other forebrain neurons than to cerebellar granule cells (fig. S14) and have larger nuclei (9 to 10 μm in diameter) (*39*, *40*), although both are similarly inactive (firing rates of 0.1 to 0.2 Hz) (*41*). It remains to be determined how granule cells in the olfactory bulb organize their genome.

More broadly, this approach showcases how life-spanning 3D genome profiling of a complex, living tissue can provide unprecedented dimensions of information. This lifelong structural transformation may point the way to previously unidentified therapeutic targets for developmental and aging-related disorders. Wide application of the 3D genome technologies developed here to many brain regions and bodily tissues may contribute to solving long-standing challenges such as dissecting the genetic basis of interindividual variability, characterizing ultrarare cell types, and revealing the full diversity and dynamics of 3D genome organization across the life of mammals.

This study has certain important limitations. For example, we used frozen human samples, which might differ from fresh samples. We also note that vDip-C does not apply to human samples; however, human Purkinje cells could alternatively be isolated by flow cytometry based on size. Future work will be required to test functional relationships between structural and transcriptional changes.

REFERENCES AND NOTES

1. Y. Chi, J. Shi, D. Xing, L. Tan, *Front. Mol. Biosci.* **9**, 959688 (2022).
2. L. Tan, D. Xing, C.-H. Chang, H. Li, X. S. Xie, *Science* **361**, 924–928 (2018).
3. L. Tan, D. Xing, N. Daley, X. S. Xie, *Nat. Struct. Mol. Biol.* **26**, 297–307 (2019).
4. L. Tan, *STAR Protoc.* **2**, 100622 (2021).
5. L. Tan *et al*., *Cell* **184**, 741–758.e17 (2021).
6. R. A. Barton, C. Venditti, *Curr. Biol.* **24**, 2440–2444 (2014).
7. F. Cauda *et al*., *J. Neurol. Neurosurg. Psychiatry* **82**, 1304–1313 (2011).
8. C.-Y. Lin, C.-H. Chen, S. E. Tom, S.-H. Kuo; Alzheimer's Disease Neuroimaging Initiative, *Cerebellum* **19**, 217–225 (2020).
9. M. J. Wagner, T. H. Kim, J. Savall, M. J. Schnitzer, L. Luo, *Nature* **544**, 96–100 (2017).
10. D. P. Hoffman *et al*., *Science* **367**, eaaz5357 (2020).
11. T. Yamada *et al*., *Nature* **569**, 708–713 (2019).
12. J. V. Goodman *et al*., *Nat. Commun.* **11**, 3419 (2020).
13. W. Tian *et al*., Epigenomic complexity of the human brain revealed by single-cell DNA methylomes and 3D genome structures. bioRxiv 2022.11.30.518285 [Preprint] (2022); https://doi.org/10.1101/2022.11.30.518285.
14. M. Sepp *et al*., Cellular development and evolution of the mammalian cerebellum. bioRxiv 2021.12.20.473443 [Preprint] (2021); https://doi.org/10.1101/2021.12.20.473443.
15. I. Sarropoulos *et al*., *Science* **373**, eabg4696 (2021).
16. S. Horvath *et al*., *Aging* **7**, 294–306 (2015).
17. L. Duncan *et al*., *Neuropsychopharmacology* **48**, 764–772 (2023).
18. V. Kozareva *et al*., *Nature* **598**, 214–219 (2021).
19. J. D. Welch *et al*., *Cell* **177**, 1873–1887.e17 (2019).
20. J. M. Granja *et al*., *Nat. Genet.* **53**, 403–411 (2021).
21. K. S. Bedi, R. Hall, C. A. Davies, J. Dobbing, *J. Comp. Neurol.* **193**, 863–870 (1980).
22. H. Heaton *et al*., *Nat. Methods* **17**, 615–620 (2020).
23. K. L. Svenson *et al*., *Genetics* **190**, 437–447 (2012).
24. B. C. Campbell *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30710–30721 (2020).
25. R. J. Platt *et al*., *Cell* **159**, 440–455 (2014).
26. Q. Huang *et al*., *PLOS ONE* **14**, e0225206 (2019).
27. K. Y. Chan *et al*., *Nat. Neurosci.* **20**, 1172–1179 (2017).
28. R. Chen *et al*., *Nat. Biotechnol.* **39**, 161–164 (2021).
29. K. Nitta, Y. Matsuzaki, A. Konno, H. Hirai, *Mol. Ther. Methods Clin. Dev.* **6**, 159–170 (2017).
30. M. G. Heffel *et al*., Epigenomic and chromosomal architectural reconfiguration in developing human frontal cortex and hippocampus. bioRxiv 2022.10.07.511350 [Preprint] (2022); https://doi.org/10.1101/2022.10.07.511350.
31. Z. Zhao *et al*., Mega-Enhancer Bodies Organize Neuronal Long Genes in the Cerebellum. bioRxiv 2023.07.19.549737 [Preprint] (2023); https://doi.org/10.1101/2023.07.19.549737.
32. E. V. Bashkirova *et al*., *eLife* **12**, RP87445 (2023).
33. K. Monahan, A. Horta, S. Lomvardas, *Nature* **565**, 448–453 (2019).
34. M. Zazhytska *et al*., *Cell* **185**, 1052–1064.e12 (2022).
35. C. Celen *et al*., *eLife* **6**, e25730 (2017).
36. Y. Takei *et al*., High-resolution spatial multi-omics reveals cell-type specific nuclear compartments. bioRxiv 2023.05.07.539762 [Preprint] (2023); https://doi.org/10.1101/2023.05.07.539762.
37. K. Powell, A. Mathy, I. Duguid, M. Häusser, *eLife* **4**, e07290 (2015).
38. M. Arancillo, J. J. White, T. Lin, T. L. Stay, R. V. Sillitoe, *J. Neurophysiol.* **113**, 578–591 (2015).
39. E. Gagyi *et al*., *Brain Pathol.* **22**, 803–810 (2012).
40. K. S. Bedi, *J. Comp. Neurol.* **311**, 425–433 (1991).
41. Y. Senzai, G. Buzsáki, *Neuron* **93**, 691–704.e5 (2017).

ACKNOWLEDGMENTS

We thank Arima for early kit access; NIMH HBCC, Stanford ADRC, and NIH NeuroBioBank for human samples; F. Zhang (MIT) and X. Jin (Scripps) for advice on *Chd8* and KASH; C. Celen (UT Southwestern) for help on *Arid1b*; H. Heaton (Auburn) for advice on souporcell; R. Corces (UCSF) for help with ArchR; D. Xing (Peking), J. Raymond (Stanford), T. Clandinin (Stanford), A. Brunet (Stanford), L. Fan (Stanford), and P. Wang (Stanford) for helpful discussion; students of SIN bootcamp (J. Bendrick, J. Cheng, C. Lewis, K. Malacon, N. Manfred, and B. Zhou) for help on experiments; Stanford PAN; and Stanford Shared FACS. We thank donors and their families, medical examiners offices, UMBTB, and SMRI brain bank for contribution to HBCC. **Funding:** BWF CASI, Baxter Foundation, Stanford MCHRI Uytengsu-Hamilton Award, Stanford Medicine Dean's Fellowship, and Stanford Berry Fellowship (L.T.); Stanford Chemistry, Stanford Bio-X, Goldwater Foundation, and Stanford Major Grant (J.S.); Stanford Barres Fellowship (B.P.); Stanford NeURO Fellowship (C.P.W.); NIH/NIA P30 AG066515 and CZI 2019-199150 (I.C.); NIMH R01 MH123486, NIMH R21 MH125358, and Stanford Jaswa Award (L.D.); NINDS K99 NS119784 (R.C.); and Gatsby Foundation and NIH (K.D.). **Author contributions:** Experiment design: L.T., J.S., I.C., L.D., R.C., and K.D.; Experiment performance: L.T., J.S., B.P., K.V., and R.C.; Data analysis: L.T., J.S., S.M., C.P.W., B.P., Y.S., K.V., I.C., L.D., R.C., and K.D.; Writing: L.T. and K.D. **Competing interests:** L.T. is an inventor on the Dip-C patent US 11,530,436 ("Multiplex end-tagging amplification of nucleic acids"). K.D. is a cofounder and a scientific advisory board member of Stellaromics and Maplight Therapeutics and a scientific advisory board member of BrightMinds Biosciences. **Data and materials availability:** Raw and processed data are available under the BioProject PRJNA933352 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA933352). Code is available from GitHub (https://github.com/tanlongzhi/dip-c). The vDip-C vector is available from Addgene (https://www.addgene.org/207613).

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh3253
Materials and Methods
Figs. S1 to S28
Tables S1 to S6
References (*42–54*)

Submitted 25 February 2023; accepted 3 August 2023
10.1126/science.adh3253

BROWN DWARFS

Evidence for a radiation belt around a brown dwarf

J. B. Climent[1,2]*, J. C. Guirado[1,3], M. Pérez-Torres[4,5], J. M. Marcaide[6,7], L. Peña-Moñino[4]

Ultracool dwarfs (UCDs) are a category of astronomical objects that includes brown dwarfs and very-low-mass stars. Radio observations of UCDs have measured their brightness as a function of time (light curves) and spectral energy distributions, providing insight into their magnetic fields. We present spatially resolved radio observations of the brown dwarf LSR J1835+3259 using very-long-baseline interferometry showing extended radio emission. The detected morphology is consistent with the presence of a radiation belt. Comparison with models indicates that the radiation belt contains energetic particles confined by magnetic mirroring. We contend that radio-emitting UCDs have dipole-ordered magnetic fields with radiation belt–like morphologies and aurorae that are similar to those of Jupiter.

Ultracool dwarfs (UCDs) are low-mass stars and substellar objects classified as having spectral types later than M6. Only a small fraction of UCDs have detectable radio emission at gigahertz frequencies (*1*). In those that do, the radio emission can be separated into a circularly polarized component that appears as bursts, which is modulated by the object's rotational period, and a quiescent, slowly varying emission component with a low degree of circular polarization. The bursting component has been linked to the electron cyclotron maser instability [ECMI (*2*)] mechanism, a coherent radio emission process that is responsible for auroral radio emission on the planets of the Solar System (*3*, *4*). However, the details of the emission mechanism may vary between UCDs and planets (*5*, *6*). The quiescent component is usually interpreted as being caused by radio emission from mildly or ultrarelativistic electrons spiraling in magnetic fields [gyrosynchrotron and synchrotron mechanisms, respectively (*7*, *8*)], which could originate in the UCD's corona and/or radiation belts (*9*–*11*). Studies of the radio emission from UCDs have analyzed their light curves (brightness as a function of time), spectral energy distributions, and dynamic spectra. Attempts to spatially resolve the radio emission using very-long-baseline interferometry (VLBI) have only detected a few objects (*12*–*14*) and found that their radio images were unresolved.

Radio observations of LSR J1835+3259

On 15 June 2021, we observed the brown dwarf (spectral type M8.5) LSR J1835+3259 (also cataloged as 2MASS J18353790+3259545) using the European VLBI Network (EVN). LSR J1835+3259 exhibits a rapid rotation [rotational period 2.84140 ± 0.00039 hours (*15*)] and is at a distance of 5.6885 ± 0.0015 parsecs (*16*) from Earth. Previous studies have shown that it has strong radio emission in both the bursting and quiescent components (*3*, *9*, *17*–*19*). Our observations resolve a complex radio morphology for this UCD (Fig. 1). The radio emission extends up to 3 milliarcseconds (mas), which corresponds to 0.017 astronomical units at our adopted distance or 33 stellar radii ($R_\star$) [for LSR J1835+3259, the $R_\star$ is 1.07 ± 0.05 Jupiter radii (R_{Jup}) (*20*)].

The light curve of our observation (Fig. 2) contains two bursts of left circularly polarized (LCP) radio emission at rotation phases ϕ_1 = 0.38 ± 0.02 and ϕ_2 = 1.33 ± 0.02, with peak flux densities of 3.7 ± 0.2 and 1.5 ± 0.2 millijansky (mJy), respectively. We interpret these as ECMI emission and refer to these two bursts as B1 and B2 hereafter. The time interval between the bursts' maxima is 2.70 ± 0.02 hours, which is close to (but different from) the rotation period [2.84140 ± 0.00039 hours (*15*)]. The right circularly polarized (RCP) emission varies slowly during the observation, reaching a maximum before the bursts occur. The LCP and RCP quiescent flux densities are very similar, as we would expect for a weakly polarized, nonbursting component of the LSR J1835+3259 radio emission.

We reconstructed images of the radio emission (*21*) of LSR J1835+3259 during burst B1 (rotation phase interval: 0.29 < ϕ <0.47; Fig. 3A). These images show that the nonbursting component (unpolarized, traced by the total intensity of radiation emitted, the Stokes *I* parameter) is spatially resolved into two radio sources separated by 2.3 ± 0.2 mas in the east-west direction, which are largely unpolarized (circular polarization <10%). By contrast, the bursting ECMI emission component (traced by circularly polarized emission, the Stokes *V* parameter) appears approximately halfway between the quiescent components. A similar morphology occurs during burst B2 (1.35 < ϕ < 1.52; Fig. 3B), although the separation of the double structure is 1.2 ± 0.2 mas, approximately half that measured during the first burst.

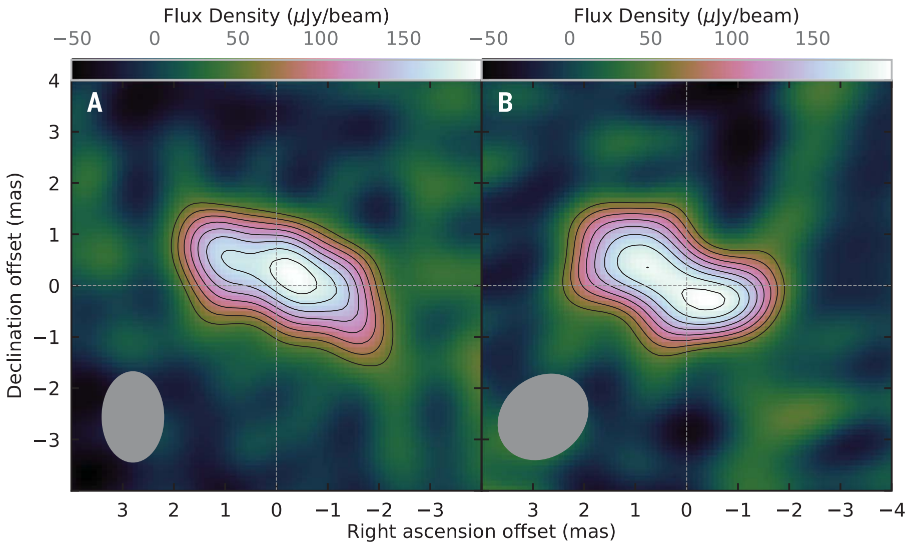


Fig. 1. Reconstructed radio images of LSR J1835+3259 using the entire observation time. (**A** and **B**) Stokes I reconstructed images of LSR J1835+3259 from the EVN observations on 15 Jun 2021 during the first rotation (A) and the second rotation (B) of the object during the observations. Contours indicate signal-to-noise ratios (S/N) of 3, 4, 5, 6, etc.; we required S/N ≥ 5 for a detection. The gray ellipses in the lower left corners indicate the full-width at half-maximum (FWHM) beam sizes, which have minor and major axes 1.14 × 1.70 mas at position angle 0.0° and 1.49 × 1.81 mas at position angle −53.9°, respectively. During each rotation, the photosphere moved −0.36 mas in right ascension and −0.19 mas in declination, which was corrected during data reduction (*21*). Both images are centered at the expected position of the optical photosphere (right ascension $18^h35^m37^s.75964$ and declination 32°59′37″.2595; indicated by the intersection of the gray lines), which was determined from the optical astrometry propagated to the observation epoch (*21*).

[1]Departament d'Astronomia i Astrofísica, Universitat de València, E-46100 Burjassot Spain. [2]Universidad Internacional de Valencia, E-46002 Valencia Spain. [3]Observatori Astronòmic, Universitat de València, Parc Científic, E-46980 Paterna Spain. [4]Instituto de Astrofísica de Andalucía, Consejo Superior de Investigaciones Científicas, E-18008 Granada Spain. [5]Facultad de Ciencias, Universidad de Zaragoza, E-50009 Zaragoza Spain. [6]Real Academia de Ciencias Exactas, Físicas y Naturales de España, E-28004 Madrid Spain. [7]Donostia International Physics Center, E-20018 San Sebastián Spain.
*Corresponding author. Email: j.bautista.climent@uv.es

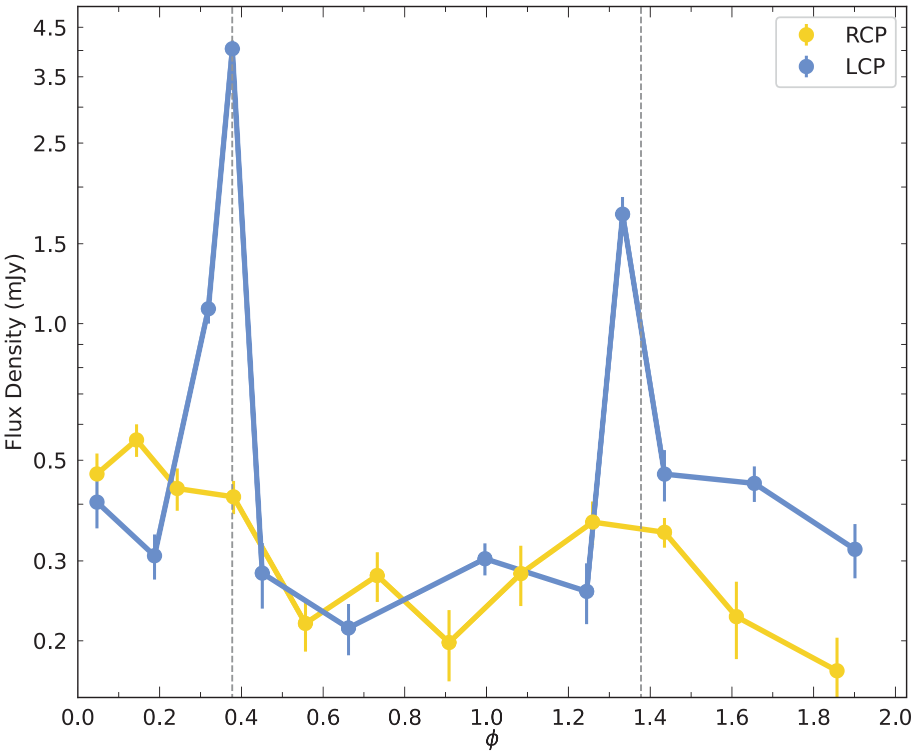


Fig. 2. Light curves of the radio emission from LSR J1835+3259 during the observations. Data points are RCP (yellow) and LCP (blue) flux densities of LSR J1835+3259, determined by integrating the selected polarization flux density over the source region (*21*) and plotted as a function of the rotation phase. The two vertical dashed lines are separated by the rotation period of the object and use the first burst as the reference phase. Error bars indicate 1σ uncertainties in the flux densities.

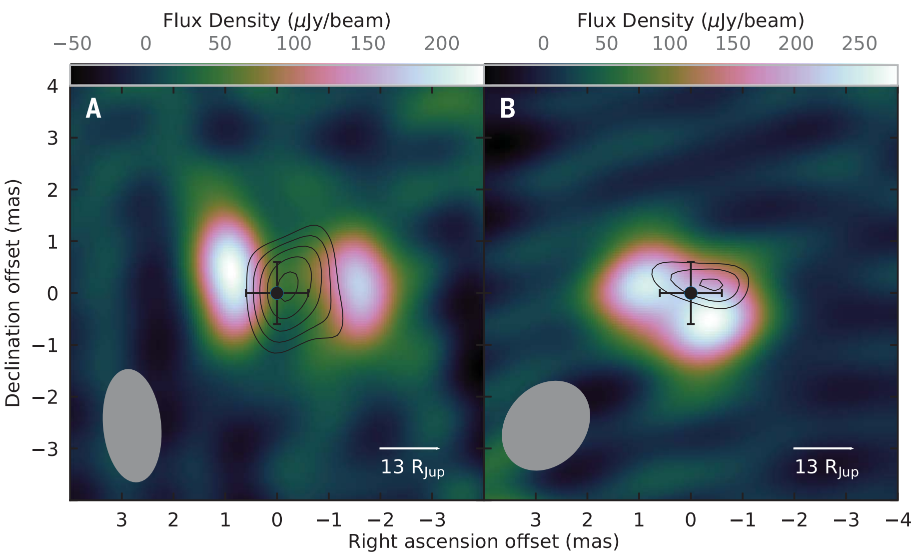


Fig. 3. Reconstructed images of LSR J1835+3259 during the radio bursts. (**A** and **B**) Same as Fig. 1, but for 30-min windows around burst B1 (A) and burst B2 (B). Ten minutes of LCP data around each burst maximum were subtracted to produce the Stokes *I* images (color). The black contours show the Stokes *V* data for each of the 10-min intervals around each burst, at S/N of 5, 6, 7, 8, and 9. The black dot indicates the expected size of the optical photosphere and its location derived from the optical astrometry, with error bars indicating the SD of our astrometric analysis (*21*).

Interpretation of the radio images

Optical astrometry of LSR J1835+3259 (*21*) gives a location coinciding with the ECMI emission (within uncertainties; Fig. 3). ECMI emission is thought to be produced near the poles of a UCD in the walls of an emission cone oriented almost perpendicular to the magnetic field lines (*22*, *23*). This implies that the magnetic axis of LSR J1835+3259 is nearly perpendicular to the line of sight, at least at the times shown in Fig. 3, A and B. We consider two possible interpretations for the quiescent double structure: (i) bipolar nonthermal emission (a magnetic axis with a position angle of ~90°) similar to that observed in other low-mass stars [such as UV Ceti (*24*)] or (ii) nonthermal emission from electrons trapped in a magnetosphere (the region around a celestial object where its magnetic field determines the motion of charged particles), forming a radiation belt (a magnetic axis with a position angle of ~0°).

In the first scenario, bipolar emission would produce different polarizations in each of the two components because they would originate in different magnetic hemispheres (*25*). Alternatively, nondipolar components could dominate the magnetic configuration of LSR J1835+3259, but we regard this as unlikely because the observed topologies of fully convective objects (*26*) indicate dipole-like arrangements for strong magnetic fields, which should include LSR J1835+3259 (*27*). There is no detectable circular polarization in the double-lobed structure in Fig. 3, A and B, so we regard the bipolar emission hypothesis as improbable.

Nonthermal radio emission from emitting structures within the magnetosphere of LSR J1835+3259 is expected to produce a double morphology if a radiation belt is seen when the magnetic equator is aligned close to the line of sight, as it does for Jupiter (*28*). For LSR J1835+3259, this would need to occur at the rotation phase of the maps shown in Fig. 3, (magnetic axis on the plane of the sky). The location of the optical emission and the absence of net circular polarization during these rotation phases provide support for the presence of a radiation belt. In this scenario, we expect synchrotron radiation to dominate the detected radio emission, as it does for Jupiter (*28*). Such a configuration would produce radio emission with linear polarization (*28*); however, our observations were only sensitive to circular polarization.

A radiation belt around LSR J1835+3259

A dipolar field constitutes a natural magnetic trap: The density of the magnetic field lines is lowest at the magnetic equator and highest at the magnetic poles. In such a configuration, charged particles are forced to spiral along the magnetic field lines, bouncing back and forth between mirror points located in each magnetic hemisphere (*29*). A radiation belt

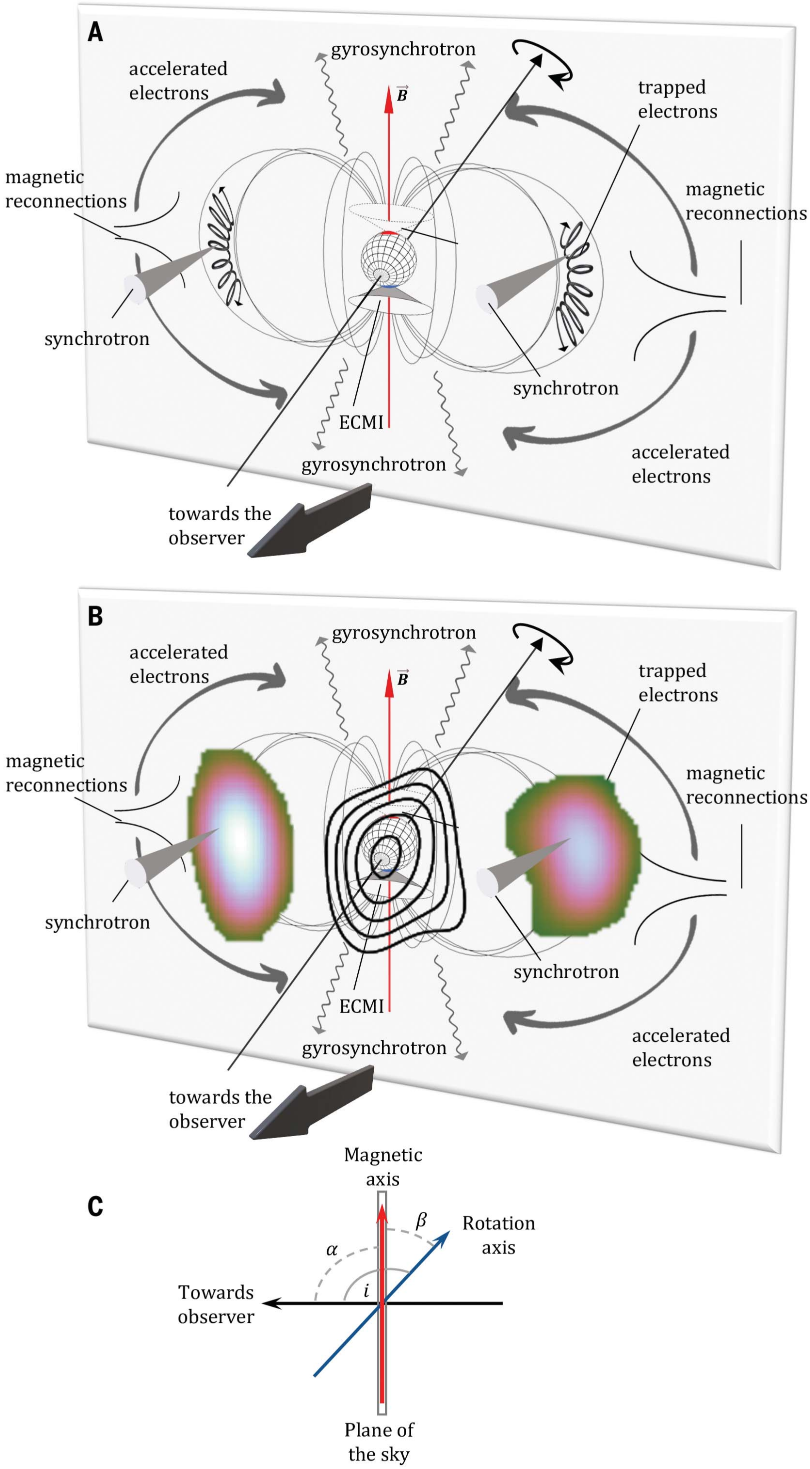


Fig. 4. Diagram of the proposed magnetosphere configuration. Shown are our proposed magnetic configurations of the magnetosphere of LSR J1835+3259 at the time of the observations (image is not to scale). (**A**) Thin black lines around the central sphere (representing the photosphere of LSR J1835+3259) indicate the UCD's magnetic field lines. Arrowed lines indicate the rotation (black) and magnetic (red) axes. Hollow cones attached to the magnetic poles (red and blue caps) represent the ECMI auroral emission. Bold curved arrows correspond to the trajectory of an accelerated electron toward the magnetic poles after magnetic reconnections, which generates gyrosynchrotron emission near the poles (wave arrows). Helical lines at each side of the magnetosphere represent the trajectory of the trapped electrons near the equator, which produce collimated synchrotron emission (narrow cones) originating at the radiation belts. (**B**) Same as (A) but with the detected radio emission during burst B1 (Fig. 3A) superimposed. The separation between the different components of the radio emission has been artificially enlarged with respect to those in Fig. 3A to more clearly indicate their locations in the model. (**C**) Compass indicating the two-dimensional orientation of the magnetic (red) and rotation (blue) axes during B1.

around LSR J1835+3259 would therefore consist of high-energy charged particles trapped in a dipolar magnetic field, similar to those present around the strongly magnetized planets of the Solar System (Earth, Jupiter, Saturn, Uranus, and Neptune). It is not known how the particles are energized even in Jupiter's radiation belts (*30*). Whatever mechanism is acting, both a fast-rotating central object and a strong magnetic field appear to be necessary to form and maintain a large and energetic radiation belt. Taking Jupiter as a reference, electrons in the magnetosphere of LSR J1835+3259 would need to be accelerated up to tens of mega–electron volts (*28*). Assuming a magnetic field in a dipole-like configuration with a strength of 5 kG in the polar regions (*27*), the magnetic field strength in such a radiation belt would be ~2 G at a distance of 13 $R_\star$ from the optical photosphere (determined from burst B1; Fig. 3A). For our observing frequency of 5 GHz, this implies an average electron energy of 21 MeV (*21*). For B2, if the same radiation belt were also responsible for the extended structure visible in Fig. 3B, then the average electron energy would need to have dropped to 8 MeV and the background magnetic field increased to 17 G. These estimates for the electron energy are similar to in situ measurements of Jupiter's radiation belts (*31*).

A radiation belt around LSR J1835+3259 with the typical energies estimated above would not be detectable during most of the object rotation if the magnetic and rotation axes were misaligned. It is therefore necessary for Fig. 3, A and B, to be produced by a specific geometric configuration, in which the magnetic equatorial plane is seen edge-on and a beam of synchrotron emission is pointing toward Earth. We estimate that a misalignment angle (β) of a few degrees would make the radiation belt undetectable (<5 σ at the observational sensitivity) in the maps before or after the bursts (*21*). Consistent with this prediction, we detect extended emission associated with the radiation belt only at rotation phases corresponding to the bursts (fig. S1), which is the most favorable orientation.

The double-lobed structure is seen only once per rotation (Fig. 3 and fig. S1). In our proposed radiation belt scenario, this constrains the inclination angle (i) and its misalignment with the magnetic axis to be i = 130 ± 10° and β = 40 ± 10°, where the uncertainties encompass the allowed range (*21*) (Fig. 4C). There are no previous determinations of i for this UCD that can be used for comparison. However, we can estimate i using a probabilistic inference method (*32*) with the rotation period (2.84140 ± 0.00039 hours), estimated radius (1.07 ± 0.05 R_{Jup}, *20*), and v sin(i) [43.9 ± 2.2 km s^{-1} (*33*)]. This yields a value of i = 120 ± 10°, consistent with the value that we estimated from the previous method.

Other sources of radio emission

Given the small beaming angle expected for synchrotron emission, the value that we estimate for β implies that synchrotron emission from the radiation belt cannot be the only radio emission mechanism during the nonbursting phases. The maximum RCP emission does not coincide with the times when the radiation belt is detected (Fig. 2). Alternatively, gyrosynchrotron emission and/or ECMI emission that is depolarized as it propagates through the magnetosphere (*19*) could be responsible for the nonbursting emission, in which case they would dominate emission at rotation phases other than those of B1 and B2 (fig. S1).

The morphological differences between Fig. 3, A and B, could be due to physical and/or instrumental effects. Physical possibilities include time variability and/or stretching of the magnetospheric field lines, as has been observed for Jupiter (*27*, *34*, *35*). Alternatively, the differing energetics of LSR J1835+3259 during the radio bursts could indicate the presence of other radiation mechanisms in addition to synchrotron (such as gyrosynchrotron and/or depolarized ECMI emission). Instrumental effects could also cause the morphological differences between B1 and B2. Given the short duration of the bursts and the cadence of observations of LSR J1835+3259 (*21*), the observations could have missed the exact rotation phase when the maximum of the LCP burst occurred.

The Stokes *V* maps in Fig. 3 show compact (milliarcsecond-scale) emission located near the optical position of LSR J1835+3259. We interpret this as auroral ECMI emission. Its detection at 5 GHz implies electron densities below 3×10^{11} cm^{-3} for the low-density regions near the poles (*21*). We regard this density as plausible for LSR J1835+3259 (*2*, *36*).

The radio powers of bursts B1 and B2 are two orders of magnitude higher than the total auroral power emitted by Jupiter (*21*), similar to previously reported bursts for this UCD (*3*). However, the profiles of B1 and B2 are substantially different (Fig. 2), which could be due to small changes in the physical parameters that determine the ECMI beam pattern (*37*). Burst B2 is 65% LCP polarized (B1 is 85%), is less intense, and has a slower decay. The energy and time decay are expected to be related: A slower decay indicates a longer trapping time of charged particles in the magnetosphere (*38*), which in turn depends on the energy of the electrons. According to this scenario, particles associated with the more energetic, fast-decaying B1 pulse quickly leave the magnetosphere, whereas those causing the less-energetic burst B2 have a longer trapping time. The different light curves of B1 and B2 (Fig. 2) could also affect the morphology of the radio emission (Fig. 3, A and B). The longer decay time of burst B2 implies that ECMI emission was more affected by scattering and depolarization within the magnetosphere of LSR J1835+3259 than was B1.

Comparison with models

We can use previous radiation belt models (*10*) to interpret the morphology and light curve of both the bursting and nonbursting emission from LSR J1835+3259. Although the models were developed for radio emission from much more massive stars (spectral types A or B), it has been proposed that they also apply to low-mass stars and gas giant planets (*10*). A diagram of our interpretation is shown in Fig. 4A. The model assumes a dipole-dominated magnetic field, with plasma permanently trapped by closed magnetic field lines, forming a radiation belt. The belt then produces, at least partially, the nonthermal radio emission (through the synchrotron mechanism, assuming similarity with Jupiter). Near the magnetic equator, electrons are accelerated by magnetic reconnection (rearrangement of magnetic field lines that releases stored energy) and propagate along the magnetic field lines, radiating nonthermal (gyrosynchrotron) emission. They eventually reach the poles of the central object, where they produce coherent pulsed ECMI auroral emission. In this interpretation, the ECMI emission is visible in the Stokes *V* map during flares B1 and B2, whereas the corresponding Stokes *I* maps (nonbursting emission coincident in time with B1 and B2) show the morphology of the radiation belt. The belt is located close to the magnetic equator, at distances up to ~13 $R_\star$ from the optical photosphere (Fig. 4B).

In the Jupiter system, the volcanic moon Io provides a source of plasma that supplies the radiation belt. It is possible that a satellite planet orbiting LSR J1835+3259 could play a similar role, as numerous rocky exoplanets have been detected around M dwarf stars [e.g., (*39, 40*)]. No companion to LSR J1835+3259 has been detected, but nor can one be ruled out (*21*). A planetary companion has previously been proposed to power the aurorae in this system (*3*). Its previously proposed parameters (*3*) would locate its orbit within the extended structure seen in our radio maps. We investigated this possibility, but can neither confirm nor refute it (see the supplementary text). If such an exoplanet exists, then it would be a potential source of plasma for the radiation belt.

REFERENCES AND NOTES

1. M. Route, *Astrophys. J.* **845**, 66 (2017).
2. G. Hallinan *et al.*, *Astrophys. J.* **653**, 690–699 (2006).
3. G. Hallinan *et al.*, *Nature* **523**, 568–571 (2015).
4. J. S. Pineda, G. Hallinan, M. M. Kao, *Astrophys. J.* **846**, 75 (2017).
5. J. Saur *et al.*, *Astrophys. J.* **859**, 74 (2018).
6. A. Gibbs, M. P. Fitzgerald, *Astron. J.* **164**, 63 (2022).
7. P. K. G. Williams *et al.*, *Astrophys. J.* **815**, 64 (2015).
8. A. G. Hughes, A. C. Boley, R. A. Osten, J. A. White, M. Leacock, *Astron. J.* **162**, 43 (2021).
9. Y. T. Metodieva *et al.*, *Mon. Not. R. Astron. Soc.* **465**, 1995–2009 (2017).
10. P. Leto *et al.*, *Mon. Not. R. Astron. Soc.* **507**, 1979–1998 (2021).
11. J. B. Climent *et al.*, *Astron. Astrophys.* **660**, A65 (2022).
12. J. Forbrich, E. Berger, *Astrophys. J.* **706**, L205–L209 (2009).
13. J. Forbrich *et al.*, *Astrophys. J.* **827**, 22 (2016).
14. Q. Zhang, G. Hallinan, W. Brisken, S. Bourke, A. Golden, *Astrophys. J.* **897**, 11 (2020).
15. P. A. Miles-Páez, S. A. Metchev, B. George, *Mon. Not. R. Astron. Soc.* **521**, 952–968 (2023).
16. Gaia Collaboration, *Astron. Astrophys.* **674**, A1 (2023).
17. E. Berger, *Astrophys. J.* **648**, 629–636 (2006).
18. E. Berger *et al.*, *Astrophys. J.* **676**, 1307–1318 (2008).
19. G. Hallinan *et al.*, *Astrophys. J.* **684**, 644–653 (2008).
20. J. C. Filippazzo *et al.*, *Astrophys. J.* **810**, 158 (2015).
21. Materials and methods are available as supplementary materials.
22. G. A. Dulk, *Annu. Rev. Astron. Astrophys.* **23**, 169–224 (1985).
23. R. A. Treumann, *Astron. Astrophys. Rev.* **13**, 229–315 (2006).
24. A. O. Benz *et al.*, *Astron. Astrophys.* **331**, 596–600 (1998).
25. R. L. Mutel, L. A. Molnar, E. B. Waltman, F. D. Ghigo, *Astrophys. J.* **507**, 371–383 (1998).
26. J. Morin *et al.*, *Mon. Not. R. Astron. Soc.* **407**, 2269–2286 (2010).
27. S. V. Berdyugina *et al.*, *Astrophys. J.* **847**, 61 (2017).
28. S. J. Bolton, R. M. Thorne, S. Bourdarie, I. DePater, B. Mauk, in *Jupiter: The Planet, Satellites and Magnetosphere*, F. Bagenal, T. E. Dowling, W. B. McKinnon, Eds. (Cambridge Univ. Press, 2004), pp. 671–688.
29. B. J. Kellett, R. Bingham, R. A. Cairns, V. Tsikoudi, *Mon. Not. R. Astron. Soc.* **329**, 102–108 (2002).
30. E. Roussos *et al.*, *Exp. Astron.* **54**, 745–789 (2022).
31. S. J. Bolton *et al.*, *Nature* **415**, 987–991 (2002).
32. K. Masuda, J. N. Winn, *Astron. J.* **159**, 81 (2020).
33. I. J. M. Crossfield, *Astron. Astrophys.* **566**, A130 (2014).
34. J. N. Girard *et al.*, *Astron. Astrophys.* **587**, A3 (2016).
35. M. F. Vogt *et al.*, *J. Geophys. Res. Space Phys.* **119**, 1925–1950 (2014).
36. M. Güdel, *Astron. Astrophys. Rev.* **12**, 71–237 (2004).

37. P. Leto *et al*.. *Mon. Not. R. Astron. Soc.* **459**, 1159–1169 (2016).
38. C. O. G. Waterfall *et al*., *Mon. Not. R. Astron. Soc.* **496**, 2715–2725 (2020).
39. G. Anglada-Escudé *et al*., *Nature* **536**, 437–440 (2016).
40. M. Gillon *et al*., *Nature* **542**, 456–460 (2017).
41. J. B. Climent, J. C. Guirado, M. Pérez-Torres, J. M. Marcaide, L. Peña-Moñino, Data and images for: Evidence for a radiation belt around a brown dwarf, Zenodo (2023); https://doi.org/10.5281/zenodo.8184164.

ACKNOWLEDGMENTS

J.B.C. thanks A. Vidotto and J. Morin for useful conversations. We thank the anonymous referees for their constructive criticisms, which greatly improved the manuscript. The European VLBI Network is a joint facility of independent European, African, Asian, and North American radio astronomy institutes. This work has made use of data from the European Space Agency (ESA) mission Gaia, processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement. **Funding:** J.B.C., J.C.G., and J.M.M. were supported by Ministerio de Ciencia e Innovación/ Agencia Estatal de Investigación (grants PGC2018-098915-B-C22 and PID2020-117404GB-C22). M.P.T. and L.P.M. were supported by Ministerio de Ciencia e Innovación (grants PID2020-117404GB-C21 and CEX2021-001131-S) and by the European Union (NextGenerationEU grant PRTR-C17.I1). J.B.C. and J.C.G. were supported by Generalitat Valenciana (grants PROMETEO/2020-080, ASFAE/2022/018, and CIPROM/2022/64). **Author contributions:** Conceptualization: J.B.C., J.C.G.; Formal analysis: J.B.C., J.C.G.; Funding acquisition: J.C.G., J.M.M., M.P.T.; Methodology: J.B.C., J.C.G., J.M.M., M.P.T.; Visualization: J.B.C., J.C.G., J.M.M., M.P.T., L.P.M.; Writing – original draft: J.B.C., J.C.G.; Writing – review and editing: J.C.G., J.B.C., J.M.M., M.P.T., L.P.M. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The raw observations are available from the EVN data archive at http://archive.jive.nl/scripts/listarch.php under observation period 2021 and experiment EC077. Our output calibrated data files and reduced images are archived at Zenodo (*41*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg6635
Materials and Methods
Supplementary Text
Figs. S1 and S2
References (*42–50*)

Submitted 12 January 2023; accepted 26 July 2023
10.1126/science.adg6635

WORKING LIFE

By **Megan Keller**

Finding my joy

"We have officially been scooped." I woke up to this dreadful news in an email from my adviser midway through the third year of my Ph.D. I was devastated. Just 6 months earlier, after 2 years of failures and dead ends, I had finally found a promising research direction to follow. Now, after dedicating all my efforts to characterizing an unknown gene with an intriguing phenotype, a preprint showed someone had beaten me to it. I began to panic. Was my Ph.D. not meant to be? Should I cut my losses and leave my program?

Desperate for perspective, I sought advice from far and wide—family, friends, classmates, students who had left my program, trusted faculty members, and former advisers. In the end I realized that no one could make the decision for me; I was the only one who could choose my future.

So, in the words of Marie Kondo, I began to look for things that sparked joy. I thought about how, ever since learning why clouds form different shapes in the fourth grade, I can't help but share fun science facts with friends and family. I remembered how watching the movie *Twilight* as a young teen had introduced me to mitosis and cellular division and spurred me to pursue molecular biology. I thought fondly back to the statewide conferences I loved attending as an undergraduate, where students presented all sorts of research and every interaction left me refueled and eager to learn more about different fields of science.



"Ever since learning why clouds form different shapes ... I can't help but share fun science facts."

I had embarked on a Ph.D. because I was excited to solve mysteries of the natural world, and I had dreams of becoming a university professor, igniting curiosity in young minds. But after reflecting on what gave me joy, I saw a new career path: informing and inspiring the public about scientific knowledge as a communicator.

I still had to decide whether to finish my Ph.D. I had no clear research project and would probably have to start from scratch. If I continued, would I be able to finish "on time"? And if I did, would that set me up for a better science communication career?

Knowing my future did not include research, my adviser and I discussed how we might scale back our ambitions for my Ph.D. project. I also sought advice from professionals working in science communication. Some said having a Ph.D. could give me an advantage with potential employers; others said all that mattered was whether I could effectively communicate complex scientific topics. Again, only I could decide.

It was a risk, but after much debate, I decided I would rather spend a few more years trying to get the Ph.D., in case it turned out to be an asset in the future. In the meantime, I sought every opportunity I could to explore science communication and develop my skills. I joined my university's newspaper, took classes and online certification programs, and completed a communications internship at a journal. It wasn't easy to convince my adviser that these were worthwhile endeavors. But eventually he came around and became my biggest supporter—and even started to work on becoming a better science communicator himself.

Two years have passed since I got scooped. I've spent long hours brainstorming for project leads and testing simple hypotheses in the lab. I've reduced the scope of the research to be publishable in smaller journals. And thanks to a project that yielded quick and promising results, I will soon finish my Ph.D., with plans to find a role in science communication after I graduate.

In conversations with my classmates, I'm finding that many have postponed addressing the question of what they will do after grad school, which often leaves them scrambling toward the end. I probably would have been in the same boat if I hadn't been scooped. So, with the benefit of hindsight, I'm grateful it happened. Losing my project forced me to reflect on my desires—and find an unexpected direction toward a brighter future. ■

Megan Keller is a Ph.D. student at Cornell University. Send your career story to SciCareerEditor@aaas.org.